In presenting the dissertation as a partial fulfillment of the requirements for an advanced degree from the Georgia Institute of Technology, I agree that the Library of the Institute shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to copy from, or to publish from, this dissertation may be granted by the professor under whose direction it was written, or, in his absence, by the Dean of the Graduate Division when such copying or publication is solely for scholarly purposes and does not involve potential financial gain. It is understood that any copying from, or publication of, this dissertation which involves potential financial gain will not be allowed without written permission.

3/17/65
b

MATHEMATICAL MODELS FOR FORECASTING HOSPITAL PERSONNEL AVAILABILITY

A THESIS

Presented to

The Faculty of the Graduate Division

by

David Richard Shaw

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Industrial Engineering

Georgia Institute of Technology

September, 1967

MATHEMATICAL MODELS FOR FORECASTING

HOSPITAL PERSONNEL AVAILABILITY

Approved:

Chairman

Date approved by Chairman: 1-15-68

TABLE OF CONTENTS

TABLE OF CONTENTS CONTINUED

ACKNOWLEDGMENTS

The direction of Dr. Harold E. Smalley was especially helpful in the preparation of this thesis. The writer is also grateful for the cooperation of the Personnel Department and the members of the Systems Development Department of the University of Alabama Hospitals and Clinics, without which much of this work could not have been accomplished.

The suggestions and assistance from Dr. J. Gordon Davis, Dr. William A. Schaffer and Mr. Walter F. Robbins were essential in this investigation.

A note of appreciation is also due Mr. Ollie B. Francis, Jr. of the Rich Electronic Computer Center for his assistance.

This investigation was supported in part by the U. S. Public Health Service of the National Institutes of Health, Division of Hospital and Medical Facilities (Grant No. HM-00430-02).

Special recognition is due the author's wife and parents and certain individuals to whom he can only express a sincere and unlimited gratitude.

LIST OF TABLES

SUMMARY

The objective of this investigation was to ascertain the mathematical relationships among selected quantitative employee characteristics. These relationships were used to identify those factors that are most highly associated with personnel turnover and absenteeism which affect the availability of hospital personnel.

The mathematical relationships developed from among the quantitative employee characteristics were used as a basis for forecasting employee length of service and absenteeism. Using the developed relationships, alternative courses of administrative action that seek to increase the availability of hospital personnel were analyzed by identifying and minimizing the effects of those factors that reduce personnel availability.

The general methodology of this investigation consisted of ascertaining the appropriate mathematical relationship that exists among the dependent variables of length of service and absenteeism rate with the selected independent variables of salary, age, dependents, previous employment and absenteeism rate. Data used in this investigation were

obtained from records of terminated employees at the University of Alabama Hospitals and Clinics for selected occupational classifications for the calendar year 1964.

The following primary conclusions were reached.

(A) With length of service as the dependent variable, multiple linear equations best fit the relationship between the dependent and independent variables for the occupational classifications of Clerk, Porter, Orderly, Maid, Licensed Practical Nurse - White, Licensed Practical Nurse - Non-White, Secretary, Technologist, Messenger and Elevator Operator. See equations (18) through (27).

(B) Independent variables of salary, age, previous employment and absenteeism were consistently found to be important in explaining the variability in employee length of service. For most occupational classifications, salary and age were the most important of the independent variables used.

(C) With absenteeism rate as the dependent variable, multiple linear equations best fit the relationship between the dependent variable and the independent variables for the occupational classifications of Technologist, Messenger and Laboratory Assistant. See equations (47) through (49).

(D) Independent variables of salary, age, dependents, previous employment and length of service were consistently found to be important in explaining the variability in employee absenteeism behavior. For most occupational classifications, salary and age were found to be the most important of the variables used.

Prediction models and the general methodology developed in this investigation provide an explicit statement of the relationships that exist among selected quantitative employee variables. Knowledge of the specific relationship for each occupational classification and the relative importance of each independent variable considered can assist hospital management to forecast not only the magnitude of turnover and absenteeism but also to evaluate alternative courses of action that seek to minimize the effects of excessive employee turnover and absenteeism.

Further investigation is recommended to evaluate the effects of existing hospital and community wage and salary structures and fringe benefit programs upon the mobility of hospital employees. Additional study is also needed to ascertain the effects that other factors such as existing transportation systems have on the availability and mobility of hospital employees.

CHAPTER I

INTRODUCTION

History and Background of the Availability of Hospital Personnel

The objective of this investigation was to ascertain the mathematical relationships among selected quantitative employee characteristics. These relationships were used to identify those factors most highly associated with personnel turnover and absenteeism which affect the availability of hospital personnel.

The mathematical relationships developed from among selected quantitative employee characteristics were used as a basis for forecasting employee length of service and absenteeism. Using the developed relationships, alternative courses of administrative action that seek to increase hospital personnel availability were analyzed by identifying and minimizing the effect of those factors that reduce hospital personnel availability.

One of the problems faced by today's industries is maintaining an efficient and effective standard of operation while undergoing a rapid growth in both size and complexity of organization. Management is often faced with the task of making decisions concerning the allocation of scarce resources. Health institutions are also faced with the problems of maintaining efficient and effective standards of operation when one considers the rapid growth and increased utilization of health facilities.

The quality of medical care received by the people of the United States is considered to be among the best in the world. Continued improvement of the quality and level of medical care is assured by such recent advances and innovations as artificial kidney units, open heart surgery and intensive care monitoring. However, the resulting costs of integrating these and other medical technological advances into today's health care system have been rapidly increasing. Expenditures for health services rose from about six billion dollars at the end of World War II to almost twenty-five billion dollars in 1964.[1] Of this expenditure almost twelve billion dollars were attributable solely to hospital care.[2] As administrative leaders of health institutions consider this rapid increase in cost and the limited availability of resources, the problems of excessive employee absenteeism and turnover must receive their increased attention.

All types of organizations throughout the nation are experiencing absenteeism and turnover problems. The average number of absent days taken per hospital employee is relatively consistent with manufacturing and service industries.[3] However, the employee turnover rate in 1957, as reported by Levine and Wright[4] is approximately 20 to 30 percent per year higher than that of other industries. The cost of turnover to hospitals was estimated to be $100,000,000 per year in 1957.

The hospital does present some special problems in that it operates seven days a week and 24 hours a day and is responsible for a "product" of extreme importance -- patient care. Care of the patient cannot be interrupted; therefore, it is often necessary to bring in replacement employees, to have employee work overtime or to have employees on standby when a shortage of personnel is experienced.

When employee absenteeism and turnover rates are high, there may be a tendency by hospital management to overstaff, especially in patient care areas, because of the concern for continuity of the desired level of care for the patients. If permitted to exist, this situation would contribute excessively to the rise in the cost of health care.

Therefore, some control over the two variables, absenteeism and turnover, would be a contribution toward controlling hospital costs.

In hospitals approximately 60 to 70 percent of the operating cost is attributable to labor costs.[5] Implementation of new welfare programs such as Medicare on July 1, 1966, to cover health care needs of a large segment of our older population and Medicaid which was designed to provide medical care for individuals of all groups who are medically indigent, have expanded the use of existing facilities. The result is an increase in costs and a need for additional personnel. On February 1, 1967, hospitals were included under the provisions of the Fair Labor Standards Act as Amended - 1966 which encompassed hospital employees under the minimum wage requirements. No longer can the hospital field curtail costs by working employees long hours at substandard wages.

Definition of the Problem of the Lack of Availability of Hospital Personnel

The critical shortage of hospital personnel in most occupational classifications is a problem recognized by many of today's health leaders.[6,7]

Francis Keppel[8] in an address to a "Conference on Job Development and Training for Workers in Health Services" in February of 1966, indicated that by 1975 there will be needed an additional one million individuals to fill the health manpower requirements of our nation. In the health services alone there will be demands for nearly ten thousand new jobs per month.

As manpower needs increase in the health field, effective utilization of manpower becomes essential. Management will need to understand and to antcipate changes in the work force such as those caused by turnover and absenteeism.

Increased recruiting programs play an important role in filling the manpower gap. Many conferences, workshops and seminars, and other programs have been designed to educate and inform the public of the advantages and opportunities of health careers. A more immediate course of action, however, is to reduce personnel turnover and the inherent reductions in productivity which can be partially attributed to the new but often untrained and inexperienced personnel.

Excessive employee absenteeism constitutes another source of loss, or reduction, in the availability of personnel. Excessive absenteeism, which is often the result of a combination of social and psychological factors, not only reduces the available manpower resource but necessitates continual staffing with associated problems of low morale and reductions in productivity. The reduction in morale and productivity are attributable, at least in part, to the uncertainties of staffing schedules which permit personnel to plan their non-hospital related activities for only a limited basis. Should personnel be flexible enough to adjust to frequent changes in staffing schedules, other problems are created when personnel are asked to work in unfamiliar areas and to perform activities other than those to which they are usually assigned. As a result, departmental staffing is frequently accomplished on a day-to-day basis with the ability to plan future staffing schedules being an exception rather than a rule.

While some recruiting activity always will be required, the importance of expending equal effort in reducing turnover and absenteeism has often been overlooked. It is plain that the seriousness of manpower shortages is reduced when turnover and absenteeism are at minimum levels, and likewise the manpower problems are aggravated when turnover and absenteeism are high. Therefore, lack of attention to patterns of absenteeism and turnover limits the possibility of a solution to these problems, and may also leave recruitment programs with inadequate guidelines for future activity. In another sense, even in the worst possible situation there may be certain groups of people, with identifiable characteristics, who consistently prove to be "better" hospital employees from the standpoint of longer lengths of service and lower absenteeism rates. Therefore, the recruitment of people with previously identified acceptable characteristics may then become more efficient.

Therefore, the identification and measurement of the effects and interrelationships of selected employee variables such as age, sex, race, salary, dependents, education, occupation, previous job experience, absenteeism and length of service, will be an important step in reducing the uncertainties encountered in explaining the causes of the variability of employee absenteeism and turnover.

Operational Environment

The University of Alabama Hospitals and Clinics is a division of the University of Alabama in Birmingham/Medical Center. This facility is a 670-bed general, acute hospital with a large outpatient department and serves as the principal teaching unit for the Medical College of Alabama, the University Hospital School of Nursing, the post-doctoral educational programs in the medical specialties and in the schools in the ancillary services. The hospital also lends support to the University of Alabama School of Dentistry and other adjunct sciences which need, on occasion, to hospitalize their patients. This is the largest general hospital in the state of Alabama and ranks among the 35 largest in the country in the amount of service rendered to patients.

The University of Alabama Hospitals and Clinics in January, 1967, had approximately 2,004 full-time people working 83 occupational classifications.

Specifically, at the University of Alabama Hospitals and Clinics, the annual turnover rate among all full-time employees ranged from 46.2 percent to 53.0 percent during the past six year period of 1960-1965. During the same six year period some occupational categories had a turnover rate as low as 17.5 percent and others as high as 101.5 percent. Absenteeism rate at the University of Alabama Hospitals and Clinics during the same six year period ranged from a low of 1.6 percent to a high of 2.4 percent per year.

CHAPTER II

LITERATURE SURVEY

Levine and Wright[9] have conducted extensive work on the magnitude, cost and methods of computation of turnover and instability rates of nursing personnel. The magnitude and cost of personnel absenteeism have been recorded by May[10] in her surveys of hospitals and other industries. Flagle[11] and Lane and Andrew[12] have conducted a number of studies dealing with the relationship of turnover and length of service. Their findings essentially revealed that the probability of an employee remaining employed increases with the length of employment. In particular, Flagle reveals that the first six months of a new employee's length of service is a critical period in relation to the employee's permanent employment. Flagle also points out that age seems to have a definite influence on the turnover of employees.

Work by Gross, Yeracoris and Grosof[13] confirms the findings of Flagle, Lane and Andrew. The work by Gross, Yeracoris and Grosof was a four year study of labor turnover in four comparable metropolitan general hospitals. The data were gathered according to organizational, demographic, social and socio-psychological characteristics. Numerous factors, such as age, marital status, length of employment and numbers of jobs in occupational history were found to be associated with labor turnover. However, neither the exact nature of the association nor a complete description of the possible interactions of factors was discussed. This study did reveal extensive consideration of the most pertinent factors that could be associated with labor turnover with the possible exception of employee absenteeism.

Gibson[14] proposes a conceptualization based upon concepts of the need-oriented individual and the goal-oriented organization linked together by contract to explain the conflicting findings on absences of personnel. Gibson also reports on a number of studies by others on absence behavior. In their 1953 study of white collar workers, Metzner and Mann[15] found that the high identification worker could be expected to be absent relatively less frequently, but when absent he can be expected to be absent for a relatively longer period of time; Shepherd and Walker[16] in their study of the relation of absenteeism to wage level and responsibility and Mann and Hoffman[17] in their study of workers and automation show that an inverse relationship exists between status of position and absence at selected levels; studies by Metzner and Mann[18] and Isambert-Jamati[19] reveal that women generally have been found to have a higher rate of absence both in frequency and duration. If, however, a woman has undertaken a career or assumes responsibility at a relatively high level, her identification might be expected to reflect a lower rate of absence; findings by Kahne, Ryder, Snegriff and Wyshak[20] in their study of older workers show that work identification becomes more positive with age and, therefore, results in less frequent absences; Kahne, Ryder, Snegriff and Wyshak[21] also show that length of service serves as an indication of the degree of identification with the organization and thus as the length of stay increases, the frequency of absences which are independent of age decrease.[22] Gibson[23] found that the higher frequency of non-illness absences, since they may be less legitimate than illness, could be expected to be associated with relatively low identification of workers. Other studies by Metzner and Mann[24] and by Shepherd and Walker[25] indicate that the rate of absences is associated with the wage level. Acton's[26] study revealed that it is not the size of the organization that makes the difference in absent rates but factors such as quality of management and supervision. Patchen's[27] study partially verifies Acton's finding that the fairness of past promotion was related to absence, independent of the prospects for future promotion and pay. Jewett's[28] investigation reveals some effect of the location of the worker's place of residence upon absence.

Several relevant investigations using more quantitative methods such as multiple regression analysis, include the study conducted by Higbie[29] to determine the relation of mental and physical characteristics to success in different occupations and the investigation by George[30] of the relation between civil service salaries and specific job characteristics. The result of these studies by Higbie and George provided a basis for measuring job success and for rating other jobs based on quantitative methods.

Regression analysis and analysis of variance are statistical techniques used in many fields of research and investigation. The use of these techniques in the fields of psychology, sociology and education have centered around the quantification of the relationship of selected personnel attributes such as test scores, intelligence quotients and physical and sociological characteristics to job and school success, crime rates and response to different types of formal education.

Recent applications of regression analysis as a forecasting model in hospital settings include that of Beenhakker[31] in predicting future hospital bed needs; Davis[32] in forecasting the demand for surgical gloves; and Sonnendecker[33] in forecasting the whole blood requirements of a hospital blood laboratory. However, the application of the methods of multiple regression analysis and analysis of variance to ascertain the quantitative relationships among selected quantitative employee variables, absenteeism and length of service, apparently has not been made, particularly in a hospital setting.

Two primary deficiencies were noted in many of the previously cited studies that have expressed the mathematical relationships among several variables. Initially there has been a lack of follow through to show which of the variables were of most importance in explaining the variance associated with length of service and absenteeism. Secondly there has been little effort exerted in simulating the variability of the dependent variable in response to changes in the dependent variables. Due to the many interrelationships among the available variables in measuring turnover and absenteeism many of the previously cited studies report that results are caused by one variable when in fact the results have not been tested independently of other variables.

The result of the literature survey initially indicates that, although efforts have been made to ascertain those variables associated with labor turnover and absenteeism, there is an apparent lack of work dealing specifically with the quantitative relationships that exist among the various employee factors. Secondly, there have been few investigations that comprehensively evaluate the interdependence among employee variables used in predicting the dependent variable.

This investigation proposes to ascertain the mathematical relationships among selected employee variables. Methods of multiple linear regression analysis will be used to ascertain the relative importance of the selected employee variables in explaining the variability found in employee absenteeism and length of service. The effects of the selected variable on employee absenteeism and turnover will be measured independent of the effects of other variables in order to isolate those variables that account for the problems. In this investigation these isolated problem areas can be used as guides in additional analyses of the effects of the variables that reduce the availability of hospital personnel.

CHAPTER III

OBJECTIVES

The purpose of this research was to ascertain the mathematical relationships among absenteeism, length of service and the quantitative employee variables of age, sex, income, number of dependents, occupational classification and previous working experience, with the aim of providing hospital management a method for describing and forecasting employee absenteeism and length of service for selected classifications of hospital employees.

Hypothesis

It was hypothesized that a mathematical relationship exists among absenteeism, length of service and other quantitative employee factors and that the resulting relationship can be used to describe and forecast personnel absenteeism and length of service.

Specific Objectives

The following specific objectives were established to ascertain the mathematical relationship among the quantitative employee variables:

1) For selected groupings of employees, length of service models were constructed by ascertaining the mathematical relationships among selected quantitative employee variables and the length of service of an employee, using multiple linear regression analyses.

2) For selected classifications of employees, absenteeism models were constructed by determining the mathematical relationships among selected quantitative employee variables and the absenteeism rate of an employee using multiple linear regression analyses.

3) Validity of the length of service and the absenteeism models were tested by forecasting results using historical data.

4) Using the length of service and absenteeism models, the results of selected courses of administrative action were forecasted. These actions included such decisions as raising salaries and hiring older people in order to increase the availability of hospital personnel by increasing the length of service and decreasing the absenteeism of employees.

Scope and Limitations

This investigation was concerned exclusively with full time employees of the University of Alabama Hospitals and Clinics, Birmingham, Alabama, who terminated during the calendar year of 1964.

All descriptive and quantitative data were obtained from information recorded in the employee's personnel file. The following variables were included: (1) occupational classification (see Categories, Table 19 in Appendix I), (2) age, (3) sex, (4) race, (5) length of employment, (6) income, (7) number of dependents, and (8) previous work experience.

Since historical data were used, the results of this investigation are most applicable to the time period when the data were gathered. Should economic and social conditions change, certain employee variables may exhibit more influence than others in describing the quantitative relationships among length of service, absenteeism and other selected employee variables. Nevertheless, the methodology followed in this investigation is sufficiently described so as to permit any changes, revisions or modifications should subsequent analyses be desired.

This research was also limited to the study of discrete and continuous quantitative employee variables such as race, sex, age, income and salary. However, it is felt that certain qualitative variables influence absenteeism and turnover in addition to thé selected quantitative employee variables. Included among these qualitative employee variables are morale, type or quality of leadership, attitude, motivation, job satisfaction and other psychological and sociological variables.

Measures of the relative worth of the qualitative employee variables in relation to their influences on absenteeism and turnover of employees are somewhat difficult to ascertain quantitatively. However, to a certain extent the influence exerted by the non-quantitative employee variables is contingent upon the quantitative employee factors considered in this study.

Employee turnover rate is frequently used as the statistical measure of the rate of termination of employees, for turnover rates are usually defined as the total number of full-time employee terminations divided by the average number of full-time people employed during a specified time period. The "turnover rate" for any one person, by necessity, must be expressed by his length of service. Consequently, the statistical analysis of the employee turnover data by methods of multiple regression analysis and analysis of variance, is most conveniently accomplished by expressing employee turnover rate data as the length of service in months per terminated employee.

Employees miss work for a number of reasons such as illness and unfavorable response to employment conditions and policies. Studies by May[34] and Groner[35] have shown that sick leave policies of hospitals, if not properly administered, provide an excellent opportunity for the employee to take "sick days" needlessly and may encourage him to do so. Therefore, any analysis of days missed from work is confounded to the extent that the variability of sick days is partially a function of the type of sick leave policy and the methods used to administer the policy. Consequently, a composite figure for absenteeism was used without differentiating between paid sick days, unpaid sick days and absent days. Vacation days, holidays and leaves of absence were not included in the determination of a composite figure for absenteeism because they are controlled primarily by hospital policy.

For comparative purposes, absenteeism is expressed as a rate; therefore, the composite absenteeism figure divided by the length of service for a specified time period would yield the desired rate.

Expressing absenteeism as a rate may impose certain limitations due to the nature of the calculation of the rate. For example, the absenteeism rate for an individual would be subject to wide variation during earlier months of employment should the number of days missed fluctuate excessively. However, as the employee's length of service increases, his rate of absenteeism tends to level off due to the cumulative effect of large numbers and, therefore, reflects less sensitivity to abrupt changes in absence behavior. The cumulative absenteeism rate may not reflect an employee's true rate of absence. The smoothing or leveling effect of the cumulative absenteeism rate does not necessarily invalidate the regression analysis approach but may impose a limited range within which it can be used with any degree of sensitivity.

Each terminated employee is classified by an occupational code from among the 83 available occupational codes. A preliminary analysis reveals that 14 of the occupational classifications -- Clerk, Porter, Orderly, Registered Nurse - White, Maid, Licensed Practical Nurse - White, Licensed Practical Nurse - Non-White, Secretary, Technologist, Messenger, Maintenance Assistant, Elevator Operator, Technician, Laboratory Assistant -- contained approximately 85 percent (84.13 percent) of the total number of terminations during the calendar year of 1964. For this reason they were selected for further analysis. A study of the employee turnover data for the years of 1960 through 1964 revealed that the absenteeism and turnover rates of full-time employees were approximately the same for each of the five years surveyed. Additional analyses revealed that the 14 previously mentioned occupational classifications consistently contained over 80 percent of the total number of terminations of full-time employees.

Most all other occupational classifications contained terminations of full-time employees during the calendar year 1964; however, they were not in sufficient numbers to lend themselves to further analysis by multiple linear regression methods.

CHAPTER IV

FORECASTING MODELS FOR THE LENGTH OF SERVICE OF HOSPITAL PERSONNEL

The objectives of the following analysis was to ascertain the mathematical relationships among selected quantitative employee factors that can be used as a means for forecasting the length of service of selected hospital personnel. This analysis consisted of the application of multiple linear regression, t-test, and analysis of variance techniques. Selected quantitative employee factors will be used as independent variables while the length of service of employees is used as the dependent variable.

Dependent Variable

Y_{Li} = Length of service (in months) of the $i^{\underline{th}}$ employee.

Independent Variables

X_{1i} = Annual salary of the $i^{\underline{th}}$ employee at the date of termination.

X_{2i} = Age (in months) of the $i^{\underline{th}}$ employee at the date of termination.

X_{3i} = Number of dependents of the $i^{\underline{th}}$ employee at the date of termination.

X_{4i} = Length of employment (in months) of the i^{th} employee on job prior to his position with the hospital.

X_{5i} = Absenteeism rate of the i^{th} employee expressed as days missed per month.

Annual salary was selected as an independent variable since it partially represents the individual's ability to achieve a specified standard of living. Numerous intangible factors such as the degree of satisfaction of work, the security and status of the position, and the opportunity to learn and advance are important; however, the achievement of sufficient monetary compensation frequently determines the degree to which such personal objectives are accomplished. Since termination, expressed as the length of service, is an immediate course of action available to employees unable to achieve desired objectives, a measurable relationship of annual salary to length of service would be useful in explaining the variability found in length of service among employees.

Based on the previously cited studies by Flagle[36] and Lane and Andrew,[37] age of an employee has also been found to be closely associated with an employee's length of service. Age at termination rather than age at hire was selected as the independent variable for this investigation because of the improved accuracy in forecasting length of service. Selected simple regression analyses comparing the relationship of age at termination and length of service with age at hire and length of service revealed significant differences in explaining the variation of an employee's length of service. Age at termination reflects the variability of an employee's length of service. Thus it is more accurate to use age at termination than age at hire in predicting an employee's length of service.

The number of dependents of an employee in part reflects his responsibilities. He must provide for them the necessities to sustain a specified standard of living. Should his present position significantly limit or hinder the attainment of his objectives, alternative courses of action frequently are followed until a satisfactory solution is achieved. The dissatisfaction with one's inability to fulfill basic needs or to achieve desired personal and family objectives is expressed in many complex and interrelated ways. Among these, change of jobs, expressed as length of service, is a readily accessible and measurable indicator. Therefore, the number of dependents was selected as a variable in explaining the variability in employee length of service.

Length of service on the job held by the individual just prior to his hospital position was chosen as another independent variable. Length of an employee's previous work experience would be useful in predicting his stability in his present job. In some occupational classifications, however, this variable would be of little importance, particularly in the case of students or individuals serving internships who are hired immediately following school graduation.

The final independent variable selected was the rate of absenteeism of an employee while the employee occupied a permanent full-time position with the hospital. This variable measures the total scheduled work days missed per month by an employee. Missed work days include such legitimate days missed as illness of the employee and illness or death in his immediate family. However, work days are missed by the employee for a combination of other reasons. These may stem from and indirectly be a measure of his dissatisfaction with his job. As reported by Kilbridge,[38] Talocchi found that a definite significant relationship exists between the level of satisfaction and absenteeism. In essence, Talocchi's finding was that as the level of satisfaction increased, absenteeism decreased.

Gibson[39] also reported numerous studies that significantly relate the absenteeism of employees to specific individual and organizational characteristics. Therefore, a significant relationship between absenteeism and length of service will be meaningful in explaining a portion of the variability of length of service among employees.

These five independent variables are continuous in nature as opposed to the discrete nature of occupational classification, race and sex which were used in addition to the quantitative variable to make selected subdivisions of data with which subsequent multiple linear regression analyses were made.

The proposed multiple regression model in equation (1) assumes that length of service of employees is a linear function of each independent variable.

$$Y_L = AX_o + BX_{1i} + CX_{2i} + DX_{3i} + EX_{4i} + FX_{5i} + \epsilon_i \qquad (1)$$

where A, B, C, D, E and F are the true regression coefficients for each related independent variable, and ϵ_i is the random error independent of the X's. For linear regression the ϵ's are assumed to be normally and independently distributed with a common variance and a mean of zero and the regression is linear.

The last element, ϵ_i, may not be truly random because it includes some factors that could be accounted for if time and costs allowed one to investigate. The least squares multiple regression for the model represented by equation (1) is given as follows:

$$\hat{Y}_{Li} = aX_o + bX_{1i} + cX_{2i} + dX_{3i} + eX_{4i} + fX_{5i} \qquad (2)$$

where $\hat{Y}_{Li} - Y_{Li} = e_i$ = the residual error for the $i^{\underline{th}}$ employee and a, b, c, d, e and f are sample estimates of the true regression coefficients.

The results from the multiple linear regression analysis, using a Burroughs 5500 computer routine for multiple regression and correlations, produced equations (3) through (16).

<u>For Clerks</u>

$$Y_L = -\ 68.5354 + 0.0249X_1 + 0.0790X_2 + 1.0538X_3 - 0.1520X_4 - 0.1582X_5 \qquad (3)$$

with the multiple coefficient of correlation R = 0.43 and a standard error of estimate of 32.38 months.

For Porters

$$Y_L = 161.0362 + 0.1111X_1 - 0.0084X_2 - 0.2382X_3 + 0.0008X_4 - 0.7918X_5 \quad (4)$$

with the multiple coefficient of correlation R = 0.49 and a standard error of estimate of 10.04 months.

For Orderlies

$$Y_L = - 177.6279 + 0.0980X_1 + 0.0973X_2 - 0.6682X_3 - 0.0863X_4 - 0.5500X_5 \quad (5)$$

with the multiple coefficient of correlation R = 0.50 and a standard error of estimate of 17.79 months.

For Registered Nurses - White

$$Y_L = + 9.0023 + 0.0002X_1 + 0.0134X_2 - 2.2080X_3 - 0.0087X_4 - 1.1879X_5 \quad (6)$$

with the multiple coefficient of correlation R = 0.21 and a standard error of estimate of 10.93 months.

For Maids

$$Y_L = 486.1876 + 0.3564X_1 + 0.1914X_2 - 6.2050X_3 - 0.0189X_4 - 1.6003X_5 \quad (7)$$

with the multiple coefficient of correlation R = 0.90 and a standard error of estimate of 20.83 months.

For Licensed Practical Nurses - White

$$Y_L = 275.0326 + 0.1143X_1 + 0.0447X_2 + 0.1385X_3 - 0.0156X_4 - 0.4615X_5 \qquad (8)$$

with the multiple coefficient of correlation R = 0.82 and a standard error of estimate of 12.73 months.

For Licensed Practical Nurses - Non-White

$$Y_L = 940.6215 + 0.3838X_1 + 0.1139X_2 + 0.2272X_3 - 0.2977X_4 + 7.0151X_5 \qquad (9)$$

with the multiple coefficient of correlation R = 0.88 and a standard error of estimate of 19.17 months.

For Secretaries

$$Y_L = -\ 31.8154 + 0.0114X_1 + 0.0337X_2 + 1.5118X_3 + 0.2088X_4 - 3.7619X_5 \qquad (10)$$

with the multiple coefficient of correlation R = 0.57 and a standard error of estimate of 16.20 months.

For Technologists

$$Y_L = +\ 34.3282 + 0.0219X_1 - 0.4545X_2 + 6.8992X_3 + 0.3429X_4 + 6.9210X_5 \tag{11}$$

with the multiple coefficient of correlation R = 0.73 and a standard error of estimate of 16.93 months.

For Messengers

$$Y_L = 239.8584 + 0.1652X_1 + 0.0196X_2 + 1.2820X_3 - 0.0635X_4 - 1.3192X_5 \tag{12}$$

with the multiple coefficient of correlation R = 0.89 and a standard error of estimate of 3.56 months.

For Maintenance Assistants

$$Y_L = +\ 39.7914 - 0.0059X_1 + 0.0661X_2 - 9.2244X_3 + 0.0036X_4 + 73.7146X_5 \tag{13}$$

with the multiple coefficient of correlation R = 0.75 and a standard error of estimate of 27.36 months.

For Elevator Operators

$$Y_L = +\ 565.7255 - 0.3706X_1 + 0.0902X_2 - 4.9212X_3 + 0.2385X_4 - 4.8298X_5 \tag{14}$$

with the multiple coefficient of correlation R = 0.94 and a standard error of estimate of 8.11 months.

For Technicians

$$Y_L = - 29.8689 - 0.0104X_1 + 0.2721X_2 - 6.3751X_3 + 0.1428X_4 - 2.9575X_5 \quad (15)$$

with the multiple coefficient of correlation R = 0.84 and a standard error of estimate of 5.55 months.

For Laboratory Assistants

$$Y_L = + 154.2664 - 0.0180X_1 - 0.3736X_2 - 9.3213X_3 - 1.4362X_4 + 17.2393X_5 \quad (16)$$

with the multiple coefficient of correlation R = 0.87 and a standard error of estimate of 6.15 months.

A partial summary of the initial multiple linear regression analysis for each of the selected occupational classifications is contained in Table 1. Other summaries of this initial analysis are found in Appendix II, Tables 20, 21, 22 and 23.

Assuming the regression to be linear, Snedecor[40] uses the analysis of variance in testing the regression of the complete equation by ascertaining the significance in the reduction in the sum of squares attributable to regression.

Therefore, the hypothesis to be tested is that no regression can be shown to be present in the population from which the sample was taken. Thus for equations (3) through (16) the hypothesis to be tested is: H_0: B = C = D = E = F = 0, i.e. all the regression coefficients are not significantly different from zero.

Calculated F-values for each equation were tested for significance at the 95 percent level of confidence with the result that for five of the occupational classifications -- Registered Nurse - White, Secretary, Maintenance Assistant, Technician and Laboratory Assistant -- the hypothesis of H_0: A = B = C = D = E = F = 0 was not rejected; therefore, the conclusion was that each of the five multiple regression equations contained regression coefficients that were not significantly different from zero, and thus no regression was found to exist in the population from which the sample was taken.

The null hypothesis for equations (3), (4), (5), (7), (8), (9), (11), (12) and (14) was rejected; this indicates that regression was found to be present in the population from which the sample was drawn. Data relevant to the computations summarized in Table 2 can be found in Appendix II, Table 24.

Table 1. Multiple Regression Coefficients of Correlation and Determination and Standard Errors for Selected Occupational Classifications

Occupational Classification	Sample Size	Coefficient of Correlation	Coefficient of Determination	Standard Error of Estimate in Months
Clerk	94	0.43	0.18	32.38
Porter	91	0.48	0.23	10.04
Orderly	62	0.49	0.24	17.78
Registered Nurse - White	44	0.21	0.04	10.93
Maid	38	0.89	0.80	20.83
Licensed Practical Nurse - White	31	0.82	0.67	12.73
Licensed Practical Nurse - Non-White	32	0.87	0.77	19.17
Secretary	30	0.56	0.32	16.19
Technologist	22	0.73	0.53	16.93
Messenger	17	0.89	0.79	3.56
Maintenance Assistant	17	0.75	0.56	27.36
Elevator Operator	12	0.94	0.88	8.12
Technician	12	0.84	0.70	5.54
Laboratory Assistant	10	0.87	0.76	6.15

Table 2. Results of a Test of Significance of Regression for Selected Occupational Classifications

Occupational Classification	N	Degrees of Freedom	Calculated F-Ratio	Table Value $\alpha = 0.05$
Clerk	94	5,88	4.00	2.33
Porter	85	5,85	5.24	2.34
Orderly	62	5,56	3.70	2.39
Registered Nurse - White	44	5,38	0.35	2.47
Maid	38	5,32	26.95	2.51
Licensed Practical Nurse - White	31	5,25	10.32	2.60
Licensed Practical Nurse - Non-White	32	5,26	17.27	2.59
Secretary	30	5,24	2.27	2.62
Technologist	22	5,16	3.65	2.85
Messenger	17	5,11	8.55	3.20
Maintenance Assistant	17	5,11	2.81	3.20
Elevator Operator	14	5,8	11.64	3.69
Technician	12	5,6	2.90	4.39
Laboratory Assistant	10	5,4	2.52	6.26

However, a more exact test used by Wine,[41] and Ezekiel and Fox[42] consists of the application of the methods of analysis of variance of a one-way classification to check the validity of the hypothesis of linearity by testing for the presence of linearity of the regression. The data of each occupational classification were assembled into groups that contained no significant variation in their independent variables. Then the sum of squares for among groups and within groups were computed for the independent variable of length of service. Separation of among group sum of squares yields the variations explained due to linearity and the deviation from linearity. Table 3 contains the results of this analysis of variance test. Data relevant to the computations summarized in Table 3 were contained in Appendix II, Table 25. The hypothesis of this test was that equations (3) through (16) were nonlinear.

The hypothesis of nonlinearity for the occupational classifications of Porter, Licensed Practical Nurse - Non-White, Licensed Practical Nurse - White, Technologist, Messenger, Maintenance Assistant, Technician and Elevator Operator failed to be rejected, i.e., there is not enough evidence to say that the regression is not linear. Insufficient data prevented the testing of the hypothesis for

Table 3. Results of Analysis of Variance Testing for Nonlinearity for Selected Occupational Classifications

Occupational Classification	Calculated F-Ratio	Table F-Ratio	Significance of Nonlinearity
Clerk	10.84	1.48	Significant
Porter	1.05	1.82	Not Significant
Orderly	26.64	2.04	Significant
Registered Nurse-White	4.03	2.90	Significant
Maid	13.25	2.47	Significant
Licensed Practical Nurse - White	0.77	3.00	Not Significant
Licensed Practical Nurse - Non-White	0.27	2.72	Not Significant
Secretary	7.62	2.77	Significant
Technologist	1.28	3.44	Not Significant
Messenger	0.51	3.48	Not Significant
Maintenance Assistant	0.87	4.39	Not Significant
Elevator Operator	0.75	5.14	Not Significant
Technician	0.11	6.94	Not Significant
Laboratory Assistant	Insufficient data for this test.		

Laboratory Assistant. The finding of nonlinearity for the occupational classifications of Clerk, Orderly and Maid is consistent with the results of test summarized in Table 2. The hypothesis of nonlinearity for the remaining occupational classifications of Registered Nurse - White and Secretary, in addition to the occupational classifications of Clerk, Orderly and Maid, was rejected and, therefore, the sample data for each classification could possibly be fitted by some higher order curve.

However, as summarized in Table 4 for all occupational classifications except Registered Nurses, for which the non-linearity hypothesis was rejected, a significant amount of the variations could be accounted for by a linear relationship when the hypothesis of linearity was tested. Data relevant to computations summarized in Table 4 are contained in Appendix II, Table 26.

Table 4. Results of Analysis of Variance Testing for Linearity for Selected Occupational Classifications

Occupational Classification	Calculated F-Ratio	Table F-Ratio	Significance of Linearity
Clerk	31.99	2.59	Significant
Porter	5.44	2.60	Significant
Orderly	66.37	2.74	Significant
Registered Nurse-White	0.99	2.30	Not Significant
Maid	223.04	3.03	Significant
Licensed Practical Nurse - White	8.79	3.48	Significant
Licensed Practical Nurse - Non-White	7.89	3.20	Significant
Secretary	9.01	3.69	Significant
Technologist	4.16	3.69	Significant
Messenger	7.50	4.26	Significant
Maintenance Assistant	0.26	4.39	Not Significant
Elevator Operator	10.92	4.39	Significant
Technician	0.19	6.26	Not Significant
Laboratory Assistant	Insufficient data for test.		

With the exception of equations (6), (13), (15) and (16) all equations can be reasonably well represented by a multiple linear regression equation. For the sake of uniformity of regression model and because of the questionable significant benefits to be obtained in ascertaining the exact higher order curves for the appropriate occupational classifications, the multiple linear regression equations will be used on the following occupational classifications: Clerk, Porter, Orderly, Maid, Licensed Practical Nurse - White, Licensed Practical Nurse - Non-White, Secretary, Technologist, Messenger and Elevator Operator.

Improved linear fits can be obtained by subdividing each occupational classification into age groups, e.g., the occupational classification of Porter was subdivided into the following five age groups: less than 21 years, 21 through 30 years, 31 through 40 years, 41 through 50 years, and 50 years and older. With a sample size of 22, the less than 21 years age group had a coefficient of correlation of 0.81 and a standard error of estimate of 5.61 months. The 21 through 30 years age group had a sample size of 19 and coefficient of correlation of 0.71 with a standard error of estimate of 14.13 months. The 31 through 40 years age group contained a sample size of 13 with a coefficient of correlation of 0.85 and a standard error of estimate of 10.31 months. The age group of 41 through 50 years had a sample size of 22 with a coefficient of correlation of 0.39 and a standard error of estimate of 3.92 months. The final age group of 51 years and older contained a sample size of 15 with a coefficient of correlation of 0.68 and a standard error of estimate of 6.65 months. Linear regression analysis for the entire occupational classification resulted in the determination of a coefficient of correlation of 0.48 with a standard error of estimate of 10.04 months. Three of the five age groups revealed a coefficient of correlation higher than the multiple coefficient of correlation for the entire occupational group. However, as reported by Ezekiel and Fox[43] with smaller sample sizes the probability of obtaining wider fluctuations in the multiple coefficient of correlation increases. Fisher[44] has provided a means of judging the probable minimum value for the correlation in the universe, with any observed value and any given size of sample. In essence, Fisher states that although one cannot be sure of the true correlation existing in the universe on the basis of the correlation shown in a given sample, one can estimate a minimum value for the true correlation, with a given probability of being wrong.

Under conditions of random sampling the graph developed by Ezekiel and Fox[45] using Fisher's methodology shows that one sample out of 20, on the average, will show a correlation coefficient with a plus and minus value as high as that observed in the sample when drawn from a universe with the stated true correlation.

Using estimates of the minimum true correlation for an equation with five independent variables, all age groups for all occupational classifications that contained a sample size greater than five were compared to the sample coefficients of correlation. Results of this analysis revealed large variations among the sample coefficients of correlation and the estimated true minimum coefficient of correlation for most age groups. Consequently with the presence of wide variability between the sample correlation coefficient and the estimated true coefficient and because the lack of sufficient sample size for uniformity of comparison among all five age groups for each occupational classification, further analysis proceeded by considering each occupational classification without subdivision by ages.

Following the selection of the linear model for each occupational classification, the t-test was used to test the significance of the regression coefficient in the prediction of Y_L for each of the equations that linear regression was found to be present. The hypotheses to be tested are that $C_i - B_k = 0$ where i = b, c, d, e and f for each of the 10 equations. If the

$$\left| t_k \right| = \frac{C_i - B_k}{S_{b_k}} \geq t\ \alpha/2;\ N\text{-}n\text{-}1, \tag{17}$$

reject the hypothesis. Where

df = N-n-1

N = number of observations

n = number of independent variables

S_{b_k} = standard error of the k^{th} regression coefficient for k = 1, 2, 3, 4, 5.

The summary of results of the t-test with confidence level of 95 percent is contained in Table 5. All significant regression coefficients in each equation in Table 5 are identified by asterisks.

Evaluating the results of the application of equation (17) results in the identification of a number of independent variables that should be eliminated from each equation due to their apparent insignificance in explaining any appreciable amount of the variability found in the dependent variable of length of service. However, due to the presence of inter-

Table 5. Results of t-test for Partial Regression Coefficients

	SALARY b_1		AGE b_2		DEPENDENTS b_3		PREVIOUS EMPLOYMENT b_4		ABSENTEEISM b_5	
Occupational Classification	Absolute Value of Calcu-lated t-Value	Tabled t-Value at $\alpha=0.05$	Absolute Value of Calcu-lated t-Value	Tabled t-Value at $\alpha=0.05$	Absolute Value of Calcu-lated t-Value	Tabled t-Value at $\alpha=0.05$	Absolute Value of Calcu-lated t-Value	Tabled t-Value at $\alpha=0.05$	Absolute Value of Calcu-lated t-Value	Tabled t-Value at $\alpha=0.05$
Clerk	2.9975*	1.991	2.9348*	1.991	0.2868	1.991	2.0486*	1.991	0.1056	1.991
Porter	4.6275*	1.994	0.8381	1.994	0.5015	1.994	0.0559	1.994	1.2347	1.994
Orderly	3.4969*	2.004	2.5702*	2.004	0.4277	2.004	2.3669*	2.004	1.0814	2.004
Maid	7.7238*	2.038	5.4102*	2.038	2.9922*	2.038	2.3892*	2.038	0.3162	2.038
Licensed Practical Nurse - White	4.9590*	2.060	2.4834*	2.060	0.0536	2.060	0.1361	2.060	2.3507*	2.060
Licensed Practical Nurse - Non-White	8.4421*	2.056	2.1956*	2.056	0.1057	2.056	1.4381	2.056	1.4914	2.056
Secretary	1.4704	2.064	0.3925	2.064	0.4714	2.064	1.4000	2.064	0.8323	2.064
Technologist	3.4515*	2.120	2.6855*	2.120	1.2311	2.120	1.1289	2.120	1.3572	2.120
Messenger	3.7065*	2.201	0.2179	2.201	0.4279	2.201	1.0818	2.201	1.2628	2.201
Elevator Operator	0.6080	2.306	3.0252*	2.306	1.3601	2.306	0.6943	2.306	1.2222	2.306

*Significant difference.

correlation among the independent variables a systematic evaluation of each net regression coefficient was necessary. In relation to the primary objectives of this investigation, forecasting length of service and absenteeism for selected classifications of employees, the significance of each net regression coefficient is not as critical as would be necessary if statements concerning the significance of the net regression coefficients were to be made. Ezekiel and Fox[46] report that if the regression equation is to be used solely as a basis for making new estimates of the value of the dependent factor to be expected for given values of the independent factors, the accuracy of the several net regression coefficients does not make such a great difference. Any deficiency in one net regression coefficient may be compensated for by an excess in another. However, this does not hold true if estimates are made for extreme values of independent variables whose regressions are subject to large errors. The reliability of the net regression coefficients is a function not only of the use to which the regression equation will be made but also of sample size and the standard error of estimate.

The statistical criterion to eliminate the independent variable whose net regression coefficient is consistently insignificant for all regression equations was used. Analysis of the results found in Table 5 revealed that the independent variable of dependents was consistently found to be nonsignificant in all equations with the exception of that for the occupational classification of Maid. Other net regression coefficients in equations for all occupational classifications were found to be insignificant; however, the true values or weights of these coefficients are difficult to evaluate unless the intercorrelation among the independent variables is of negligible size, or one or more of the intercorrelated variables are eliminated.

These unwanted intercorrelations are suppressed in controlled experiments by using "balance designs" where such predictors are controlled so that they are orthogonal to each other, i.e., the intercorrelation is made to zero. Significant intercorrelation among the independent variables was not found; therefore, further analysis proceeded without having to resort to other experimental designs.

Following the elimination of dependents as an independent variable, a four variable linear regression equation was obtained for each occupational classification. Applying equation (17) the analysis of the significance of the net regression coefficients is summarized in Table 6. Other

Table 6. Results of t-test for Partial Regression Coefficients

Occupational Classification	SALARY b_1		AGE b_2		PREVIOUS EMPLOYMENT b_3		ABSENTEEISM b_4	
	Absolute Value of Calculated t-Value	Tabled t-Value at α=0.05	Absolute Value of Calculated t-Value	Tabled t-Value at α=0.05	Absolute Value of Calculated t-Value	Tabled t-Value at α=0.05	Absolute Value of Calculated t-Value	Tabled t-Value at α=0.05
Clerk	3.0067*	1.990	3.2371*	1.990	2.0658*	1.990	0.0548	1.990
Porter	4.6242*	1.993	0.9531	1.993	0.0347	1.993	1.2302	1.993
Orderly	3.5002*	2.002	2.6147*	2.002	2.3877*	2.002	1.0767	2.002
Maid	7.2756*	2.036	4.0527*	2.036	2.2602*	2.036	0.0441	2.036
Licensed Practical Nurse - White	6.1181*	2.056	2.5321*	2.056	0.1510	2.056	2.4052*	2.056
Licensed Practical Nurse - Non-White	8.7471*	2.052	2.4987*	2.052	1.4769	2.052	1.5965	2.052
Secretary	1.8860	2.060	0.6440	2.060	1.3394	2.060	0.7895	2.060
Technologist	3.3418*	2.110	2.3513*	2.110	0.8360	2.110	1.3283	2.110
Messenger	4.9603*	2.179	1.3668	2.179	1.3369	2.179	2.4715*	2.179
Elevator Operator	0.0483	2.262	3.4886*	2.262	0.0264	2.262	0.9239	2.262

*Significant difference.

results obtained from the regression analysis following the removal of the independent variable, number of dependents, are summarized in Appendix II, Table 27. As shown in Table 6, none of the four remaining independent variables were consistently insignificant; therefore, following the removal of the independent variables which do not contribute significantly to the prediction of Y_{Li}, the resulting equations are as follows:

For Clerks

$$Y_L = -67.9510 + 0.0245X_1 + 0.0817X_2 - 0.1524X_3 - 0.0804X_4 \tag{18}$$

with the multiple coefficient of correlation R = 0.43 and a standard error of estimate of 32.21 months.

For Porters

$$Y_L = -159.7714 + 0.1103X_1 - 0.0093X_2 - 0.0005X_3 - 0.7853X_4 \tag{19}$$

with the multiple coefficient of correlation R = 0.48 and a standard error of estimate of 10.00 months.

For Orderlies

$$Y_L = -\ 174.8983 + 0.0971X_1 + 0.0912X_2 - 0.0864X_3 - 0.5434X_4 \tag{20}$$

with the multiple coefficient of correlation R = 0.49 and a standard error of estimate of 17.66 months.

For Maids

$$Y_L = -\ 496.9486 + 0.3717X_1 + 0.1371X_2 - 0.1995X_3 - 0.2477X_4 \tag{21}$$

with the multiple coefficient of correlation R = 0.87 and a standard error of estimate of 23.21 months.

For Licensed Practical Nurses - White

$$Y_L = -\ 276.6025 + 0.1149X_1 + 0.0446X_2 - 0.0167X_3 - 0.4621X_4 \tag{22}$$

with the multiple coefficient of correlation R = 0.82 and a standard error of estimate of 12.48 months.

For Licensed Practical Nurses - Non-White

$$Y_L = -\ 938.8269 + 0.3828X_1 + 0.1161X_2 - 0.2929X_3 + 7.1387X_4 \tag{23}$$

with the multiple coefficient of correlation R = 0.87 and a standard error of estimate of 18.82 months.

For Secretaries

$$Y_L = -\ 40.5834 + 0.0129X_1 + 0.0498X_2 + 0.1844X_3 - 3.4818X_4 \tag{24}$$

with the multiple coefficient of correlation R = 0.56 and a standard error of estimate of 15.94 months.

For Technologists

$$Y_L = +\ 12.4416 + 0.0214X_1 - 0.3592X_2 + 0.2496X_3 + 6.8749X_4 \tag{25}$$

with the multiple coefficient of correlation R = 0.70 and a standard error of estimate of 17.18 months.

For Messengers

$$Y_L = -\ 262.6905 + 0.1759X_1 + 0.0540X_2 - 0.0716X_3 - 1.6536X_4 \tag{26}$$

with the multiple coefficient of correlation R = 0.89 and a standard error of estimate of 3.44 months.

For Elevator Operators

$$Y_L = -\ 59.1588 + 0.0270X_1 + 0.1031X_2 - 0.0081X_3 - 3.7400X_4 \tag{27}$$

with the multiple coefficient of correlation R = 0.92 and a standard error of estimate of 8.49 months.

As reported by Duncan[47] the 0.95 percent confidence limits for universe regression coefficients are given by the sample coefficient plus and minus $t_{0.025}$ times the estimated standard error of the coefficient, and the equation for 0.95 percent confidence limits for the regression value is

$$Y_{Li} = \overline{X}_i + b_{12.3...k}X_2 + \dots + b_{1k.23...k-1}X_k \pm t_{0.025}S_{1.23...k} \sqrt{1/N + C_{22}X_2{}^2 + \dots + C_{kk}X_k{}^2 + 2C_{23}X_2X_3 + \dots + 2C_{2k}X_2X_k + \dots + 2C_{3k}X_3X_k + \dots} \tag{28}$$

where $t_{0.025}$ is that for n = N - k, and k = the number of total variables. If 1 is added to the expression under the radical, an expression for the 0.95 prediction limits for an individual value is obtained. The confidence limits for multiple regression values are a pair of parabolic surfaces on either side of the sample plane of regression.

Equation (28) can be written as:

$$Y_L = a + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 \pm t_{0.025} \tag{29}$$

$$\sqrt{S_e^2 \left[1 + 1/N + \Sigma\Sigma\, C_{ij}\ (X_i - \overline{X}_i)\ (X_j - \overline{X}_j)\right]}$$

with origin at (o, $\overline{X}_1$, $\overline{X}_2$, $\overline{X}_3$, $\overline{X}_4$) for the 0.95 percent confidence limits for individual values for a multiple regression equation containing four independent variables. For example, the limiting loci for randomly picked single observations for each occupational classification are shown in Table 7. Data used to calculate the predicted length of service values (Y_{Li}) were summarized in Table 28 in Appendix II.

The 95 percent confidence limits for single observations for the multiple linear regression equation states that in the long run less than 5 percent of the observations will fall outside the 95 percent confidence limit. In each occupational classification a better estimate of Y_{Li} will be made as each independent variable X_i approaches the mean variable $\overline{X}_i$.

The general methodology followed in computing the confidence limits is presented in Table 7. The procedure for deriving the C_{ij} values and a listing of the C_{ij} values

Table 7. 95 Percent Confidence Limits for Length of Service for Individual Values of the Independent Variables

Occupational Classification	Actual Length of Service (Months)	Predicted (Y_{Li}) Length of Service (Months)	Residual Error (Months)	Confidence Limits
Clerk	2	10.9	8.9	0,22.3
Porter	11	9.8	1.2	0,29.6
Orderly	2	19.9	17.9	11.3,28.5
Maid	21	9.8	11.2	0,19.7
Licensed Practical Nurse-White	11	10.9	0.1	3.3,18.5
Licensed Practical Nurse-Non-White	51	77.9	26.9	68.6,87.2
Secretary	29	7.3	21.7	0,15.9
Technologist	12	11.2	9.8	2.0,20.4
Messenger	14	18.4	4.4	8.0,28.8
Elevator Operator	4	10.1	6.1	0,30.9

appears in Appendix II.

The 0.95 percent confidence limits for sample mean values is

$$Y_L = Y \pm t_{0.025}\sqrt{S_e^2 \left[1/M + 1/N + \sum\sum(X_i-\bar{X}_i)\ (X_j-\bar{X}_j)\right]} \quad (30)$$

where M = the sample size.

Therefore, the mean values of the independent variables for a given sample of employees possessing similar characteristics can be used in predicting the mean length of stay for these employees with 95 percent confidence limits calculated using equation (30).

Relative importance of each selected independent variable in explaining the variability associated with hospital employee length of service was ascertained by calculating standard partial regression coefficients.

Standard partial regression coefficients denoted by 1_{b_i} are the partial regression coefficients when each variable is in standard measure, that is, a deviation from the mean in units of its standard deviation. Since each 1_b is independent of the original units of measurement, a comparison of any two indicates the relative importance of the independent variable involved in predicting the dependent variable. The standard partial regression coefficient is obtained by the following equation:

$$1_{b_i} = b_i \frac{S_i}{S_y} \tag{31}$$

Table 29 found in Appendix II, summarizes the calculated standard partial regression coefficient for each occupational classification.

For each occupational classification the standard partial regression coefficient was divided by the minimum coefficient value to clearly indicate the relative importance of the independent variables. Relative rank of independent variables are summarized in Table 8.

Table 8. Relative Rank of Independent Variables with Length of Service as the Dependent Variable

Occupational Classification	Salary $^1b_1$	Age $^1b_2$	Previous Employment $^1b_4$	Absenteeism Rate $^1b_5$
Clerk	55.6	63.8	41.8	1.0
Porter	1.0	31.0	1.0	33.7
Orderly	3.5	2.4	4.1	1.0
Maid	1821.8	1106.7	621.0	1.0
Licensed Practical Nurse - White	3.9	1.8	1.0	1.6
Licensed Practical Nurse - Non-White	5.8	1.2	1.0	1.1
Secretary	1.0	3.6	8.4	403.3
Technologist	2.4	3.0	1.0	96.6
Messenger	3.8	1.0	1.0	1.9
Elevator Operator	1.8	18.2	1.0	2.7

Wide variation exists among the independent variables; however, salary and age are consistently ranked as most important for eight of the ten occupational classifications. Size of the relative values have meaning only within occupational classifications and not among classifications.

CHAPTER V

FORECASTING MODELS FOR THE ABSENTEEISM RATE OF A HOSPITAL EMPLOYEE

The goal of this analysis is to develop a technique for ascertaining the mathematical relationship among the absenteeism rate of hospital personnel and selected quantitative employee factors. Using the methodology similar to that followed in Chapter IV, this analysis will consist of the application of multiple linear regression analyses, t-tests, and analysis of variance techniques. Similar to Chapter IV, this chapter will use selected quantitative employee factors as independent variables.

Dependent Variable

Y_{Ai} = Absenteeism Rate (days per month) for the i^{th} employee.

Independent Variables

X_{1i} = Annual salary of the i^{th} employee at the date of termination.

X_{2i} = Age (in months) of the i^{th} employee at the date of termination.

X_{3i} = Number of dependents of the i^{th} employee at the date of termination.

X_{4i} = Length of employment (in months) of the $i^{\underline{th}}$ employee on job prior to their position with the hospital.

X_{5i} = Length of service (in months) of the $i^{\underline{th}}$ employee.

Salary, age, dependents and length of service on job held prior to the position at the hospital were selected as independent variables on a premise similar to that stated in Chapter IV. Job dissatisfaction, whether for economic, social, professional growth and development or any combination of reasons, is frequently expressed and partially measured not only by length of service of employees but also by the employee's absence behavior. Consequently the selection of the variables of salary, age, number of dependents, length of service of previous employment were selected as independent variables in order to ascertain the amount of variability in absenteeism rate explained by these variables.

As expressed previously, the length of service of an employee with the hospital is a partial measure of the job satisfaction and, therefore, will be useful as an independent variable in explaining the variability found in employee absenteeism rates.

The multiple regression model proposed in equation (1) in Chapter IV will be modified by replacing the independent variable of absenteeism rate with the dependent variable of length of service. The resulting proposed multiple regression model is:

$$Y_A = AX_o + BX_{1i} + CX_{2i} + DX_{3i} + EX_{4i} + FX_{5i} + \epsilon_i \qquad (32)$$

where A, B, C, D, E and F are the true regression coefficients for each related independent variable, and ϵ_i is the random error independent of the X's. As in the discussion in Chapter IV, it was assumed that the ϵ_i's are normally and independently distributed with a zero mean, a common variance and that regression is linear.

The resulting least squares multiple regression for the model represented by equation is given as follows:

$$\hat{Y}_{Ai} = aX_o + bX_{1i} + cX_{2i} + dX_{3i} + eX_{4i} + fX_{5i} \qquad (33)$$

Where $\hat{Y}_{Ai} - Y_{Ai} = e_i$ = the residual error for the $\underline{i^{th}}$ employee and a, b, c, d, e and g are sample estimates of the true regression coefficients.

The results from the multiple linear regression analysis produced equations (34) through (47).

For Clerks

$$Y_A = 1.9237 - 0.0003X_1 - 0.0018X_2 - 0.4453X_3 + 0.0078X_4 - 0.0008X_5 \quad (34)$$

with the multiple coefficient of correlation R = 0.25 and a standard error of estimate of 2.30 days per month.

For Porters

$$Y_A = -0.2343 + 0.0014X_1 - 0.0024X_2 - 0.0201X_3 + 0.0041X_4 - 0.0223X_5 \quad (35)$$

with the multiple coefficient of correlation R = 0.24 and a standard error of estimate of 1.68 days per month.

For Orderlies

$$Y_A = -3.4894 + 0.0014X_1 + 0.0127X_2 - 0.1158X_3 - 0.0100X_4 - 0.0372X_5 \quad (36)$$

with the multiple coefficient of correlation R = 0.20 and a standard error of estimate of 4.61 days per month.

For Registered Nurses - White

$$Y_A = -1.4774 + 0.0003X_1 + 0.0059X_2 - 0.2028X_3 - 0.0137X_4 - 0.0198X_5 \quad (37)$$

with the multiple coefficient of correlation R = 0.40 and a standard error of estimate of 1.41 days per month.

For Maids

$$Y_A = + 0.2161 + 0.0003X_1 + 0.0004X_2 - 0.0486X_3 + 0.0029X_4 - 0.0019X_5 \quad (38)$$

with the multiple coefficient of correlation R = 0.25 and a standard error of estimate of 0.73 days per month.

For Licensed Practical Nurses - White

$$Y_A = - 105.9568 + 0.0418X_1 + 0.0374X_2 - 0.6152X_3 - 0.0227X_4 - 0.3923X_5 \quad (39)$$

with the multiple coefficient of correlation R = 0.50 and a standard error of estimate of 11.74 days per month.

For Licensed Practical Nurses - Non-White

$$Y_A = + 13.6474 - 0.0052X_1 - 0.0015X_2 + 0.0102X_3 - 0.0048X_4 + 0.0112X_5 \quad (40)$$

with the multiple coefficient of correlation R = 0.42 and a standard error of estimate of 0.77 days per month.

For Secretaries

$$Y_A = +\ 2.5949 - 0.0002X_1 - 0.0055X_2 + 0.1019X_3 + 0.0149X_4 - 0.0075X_5 \tag{41}$$

with the multiple coefficient of correlation R = 0.46 and a standard error of estimate of 0.72 days per month.

For Technologists

$$Y_A = +\ 0.2742 - 0.0009X_1 + 0.0168X_2 - 0.1101X_3 - 0.0170X_4 + 0.0149X_5 \tag{42}$$

with the multiple coefficient of correlation R = 0.64 and a standard error of estimate of 0.79 days per month.

For Messengers

$$Y_A = -\ 49.1046 + 0.0265X_1 + 0.0552X_2 - 1.7509X_3 - 0.0240X_4 - 0.0960X_5 \tag{43}$$

with the multiple coefficient of correlation R = 0.81 and standard error of estimate of 0.96 days per month.

For Maintenance Assistants

$$Y_A = -\ 0.4370 + 0.0001X_1 - 0.0000X_2 + 0.0101X_3 - 0.0019X_4 + 0.0051X_5 \tag{44}$$

with the multiple coefficient of correlation R = 0.77 and a standard error of estimate of 0.59 days per month.

For Elevator Operators

$$Y_A = + 70.6912 - 0.0450X_1 + 0.0016X_2 - 0.3168X_3 + 0.0267X_4 - 0.0326X_5 \quad (45)$$

with the multiple coefficient of correlation R = 0.52 and a standard error of estimate of 0.67 days per month.

For Technicians

$$Y_A = - 0.6933 - 0.0008X_1 + 0.0120X_2 - 0.5539X_3 + 0.0282X_4 - 0.0133X_5 \quad (46)$$

with the multiple coefficient of correlation R = 0.78 and a standard error of estimate of 0.37 days per month.

For Laboratory Assistants

$$Y_A = - 10.3103 + 0.0009X_1 - 0.0295X_2 + 0.2329X_3 + 0.1014X_4 + 0.0219X_5 \quad (47)$$

with the multiple coefficient of correlation R = 0.99 and a standard error of estimate of 0.22 days per month.

A summary of the initial multiple linear regression analysis for each of the selected occupational classifications is contained in Table 9.

Table 9. Multiple Regression Coefficients of Correlation and Determination and Standard Errors for Selected Occupational Classifications

Occupational Classification	Sample Size	Coefficient of Correlation	Coefficient of Determination	Standard Error of Estimate in Months
Clerk	94	0.25	0.06	2.30
Porter	91	0.24	0.06	1.68
Orderly	62	0.20	0.04	4.63
Registered Nurse - White	44	0.40	0.16	1.41
Maid	38	0.25	0.06	0.73
Licensed Practical Nurse - White	31	0.50	0.25	11.74
Licensed Practical Nurse - Non-White	32	0.42	0.18	0.77
Secretary	30	0.46	0.21	0.72
Technologist	22	0.64	0.41	0.79
Messenger	17	0.81	0.65	0.96
Maintenance Assistant	17	0.77	0.59	0.23
Elevator Operator	14	0.52	0.27	0.67
Technician	12	0.78	0.60	0.37
Laboratory Assistant	10	0.99	0.98	0.22

Analysis of variance will be used in a manner similar to that in Chapter IV. The regression of the complete equation will be tested by ascertaining the significance in the reduction in the sum of squares attributable to regression. Therefore, the hypothesis to be tested is that no regression can be shown to be present in the population from which the sample was taken. Thus for equations (34) through (47) the hypothesis is: H_o: $B = C = D = E = F = 0$, i.e. all the regression coefficients are not significantly different from zero. A summary of the results is contained in Table 10.

Assuming the regression to be linear the F-value calculated from the one way analysis of variance for each equation was tested for significance at the 95 percent level of confidence. For 12 of the 14 occupational classifications -- Clerk, Porter, Orderly, Registered Nurse - White, Maid, Licensed Practical Nurse - White, Licensed Practical Nurse - Non-White, Secretary, Technologist, Maintenance Assistant, Elevator Operator and Technician -- the hypothesis of H_o: $A = B = C = D = E = F = 0$ was not rejected and, therefore, the conclusion was that each of the 12 multiple regression equations contained regression coefficients that were not significantly different from zero. Thus no regression was found to exist in the population from which the sample was taken.

Table 10. Results of a Test of Significance of Regression for Selected Occupational Classifications

Occupational Classification	Degrees of Freedom	Calculated F-Ratio	Table Value $\alpha = 0.05$
Clerk	5,88	1.14	2.33
Porter	5,85	1.02	2.34
Orderly	5,56	0.45	2.39
Registered Nurse - White	5,38	1.45	2.47
Maid	5,32	0.43	2.51
Licensed Practical Nurse - White	5,25	1.63	2.60
Licensed Practical Nurse - Non-White	5,26	1.11	2.59
Secretary	5,24	1.27	2.62
Technologist	5,16	2.19	2.85
Messenger	5,11	4.13	3.20*
Maintenance Assistant	5,11	3.11	3.20
Elevator Operator	5,8	0.58	3.69
Technician	5,6	1.81	4.39
Laboratory Assistant	5,4	51.42	6.26*

*Significant difference.

The null hypothesis for the remaining occupational classifications of Messenger and Laboratory Assistant equations was rejected indicating that regression was found to be present in the population from which the sample was drawn. Data relevant to computations summarized in Table 10 can be found in Table 30, Appendix III.

To check the validity of the assumption of linearity the data for each occupational classification were grouped in such a way that no significant variation existed among independent variables. From this grouping the among groups and within groups sum of squares are computed for the dependent variable of absenteeism. Division of the among sum of squares yields the variation explained due to linearity and the deviation from linearity. Table 11 contains the results of the one way analysis of variance testing the null hypothesis of nonlinearity.

Results of the analysis of variance summarized in Table 11 indicate that there is not sufficient evidence to accept the hypothesis of nonlinearity for all occupational classifications with the exception of Technologist. The occupational classifications of Clerk, Porter, Orderly, Registered Nurse - White, Maid and Licensed Practical Nurse - White contained sufficient variation in the five independent

Table 11. Results of Analysis of Variance Testing for Nonlinearity for Selected Occupational Classifications

Occupational Classification	Calculated F-Ratio	Table F-Ratio	Significance of Nonlinearity
Clerk	Insufficient data		
Porter	Insufficient data		
Orderly	Insufficient data		
Registered Nurse - White	Insufficient data		
Maid	Insufficient data		
Licensed Practical Nurse - White	Insufficient data		
Licensed Practical Nurse-Non-White	1.11	3.18	Not Significant
Secretary	0.80	3.20	Not Significant
Technologist	14.50	3.69	Significant
Messenger	1.64	4.39	Not Significant
Maintenance Assistant	1.00	8.85	Not Significant
Elevator Operator	1.62	6.39	Not Significant
Technician	8.00	9.28	Not Significant
Laboratory Assistant	0.17	10.13	Not Significant

variables that the one way analysis of variance was not applicable.

Results of the test of the hypothesis of linearity are contained in Table 12 for those occupational classifications that contain sufficient data for performing a one way analysis of variance. Data relevant to the computations summarized in Table 12 are contained in Table 31, Appendix III.

Table 12. Results of Analysis of Variance Testing for Linearity for Selected Occupational Classifications

Occupational Classification	Calculated F-Ratio	Table F-Ratio	Significance of Linearity
Licensed Practical Nurse-Non-White	1.78	3.69	Not Significant
Secretary	0.96	3.69	Not Significant
Technologist	17.13	3.69	Significant
Messenger	5.32	4.39	Significant
Maintenance Assistant	3.20	9.01	Not Significant
Elevator Operator	0.76	6.26	Not Significant
Technician	8.33	9.01	Not Significant
Laboratory Assistant	42.33	9.01	Significant

For the occupational classifications of Technologist, Messenger and Laboratory Assistant the null hypothesis was rejected and it was concluded that there was not enough evidence to say the relationship is nonlinear. A conclusion of linearity for Messenger and Laboratory Assistant is consistent with the finding summarized previously in Table 10. The failure to reject the null hypothesis of linearity for the remaining occupational classifications -- Licensed Practical Nurse - Non-White, Secretary, Maintenance Assistant, Elevator Operator and Technician -- is also consistent with the findings previously summarized in Table 10. That is, no relationship was found to be present in the population from which the sample was taken. Consequently, further analysis is devoted to the occupational classifications of Technologist, Messenger and Laboratory Assistant since each is best represented by a linear relationship. Analysis of the desirability of subdividing these three occupational classifications resulted in a conclusion similar to that discussed in Chapter IV, i.e., a division of each occupational classification into age categories resulted in insufficient data in any one age group to permit significant statistical inferences.

The t-test was used to determine the significance of the regression coefficient in the prediction of Y_A for each of the three equations that linear regression was found to be present. The hypotheses to be tested are that $C_i - B_k = 0$ where i = b, c, d, e and f for each of the three equations. Using equation (17) from Chapter IV, a summary of the results of the t-tests with $\alpha = 0.05$ is found in Table 13. All significant regression coefficients in each equation are identified by asterisks.

The criterion used in Chapter IV for the elimination of insignificant variables, stated that the independent variable consistently found to be insignificant when applying the t-test to the partial regression coefficient, would be omitted.

Table 13. Results of t-test for Partial Regression Coefficients

	SALARY b_1		AGE b_2		DEPENDENTS b_3		PREVIOUS EMPLOYMENT b_4		LENGTH OF SERVICE b_5	
Occupational Classification	Absolute Value of Calcu-lated t-Value	Tabled t-Value at $\alpha=0.05$	Absolute Value of Calcu-lated t-Value	Tabled t-Value at $\alpha=0.05$	Absolute Value of Calcu-lated t-Value	Tabled t-Value at $\alpha=0.05$	Absolute Value of Calcu-lated t-Value	Tabled t-Value at $\alpha=0.05$	Absolute Value of Calcu-lated t-Value	Tabled t-Value at $\alpha=0.05$
Technologist	2.9575*	2.120	1.9774	2.120	0.4067	2.120	1.2108	2.120	1.3572	2.120
Messenger	1.6386	2.201	3.1054*	2.201	2.8211*	2.201	1.6036	2.201	1.2628	2.201
Laboratory Assistant	5.0114*	2.776	7.4386*	2.776	0.9986	2.776	5.5141*	2.776	1.5588	2.776

*Significant difference.

Applying this criterion to the partial regression coefficient of the multiple linear equations for the occupational classifications of Technologist, Messenger and Laboratory Assistant, the independent variable of length of service was found to be consistently insignificant. Since the primary objective of this investigation is accuracy of forecasting rather than significance of each partial regression coefficient, the independent variable of length of service was retained. Additional justification for retaining this variable was the relative ease with which this information could be obtained.

Therefore, the resulting multiple linear equations used to represent the functional relationship between the absenteeism rate and other quantitative employee variables are stated as follows:

<u>For Technologists</u>

$$Y_A = 0.2742 - 0.0009X_1 + 0.0168X_2 - 0.1101X_3 - 0.0170X_4 + 0.0149X_5 \qquad (48)$$

with the multiple coefficient of correlation R = 0.64 and a standard error of estimate of 0.79 days per month.

<u>For Messengers</u>

$$Y_A = -49.1046 + 0.0265X_1 + 0.0552X_2 - 1.7509X_3 - 0.0240X_4 - 0.0960X_5 \quad (49)$$

with the multiple coefficient of correlation R = 0.81 and a standard error of estimate of 0.96 days per month.

<u>For Laboratory Assistants</u>

$$Y_A = -10.3103 + 0.0009X_1 - 0.0295X_2 + 0.2329X_3 + 0.1014X_4 + 0.0219X_5 \quad (50)$$

with the multiple coefficient of correlation R = 0.99 and a standard error of estimate of 0.22 days per month.

The 95 percent confidence limits can be ascertained by applying equation (29) for randomly selected single values. For example the limiting loci for randomly picked single observations for each occupational classification are shown in Table 14.

The 95 percent confidence limits for single observations for the multiple linear regression equation states that in the long run less than 5 percent of the observations will fall outside the 95 percent confidence limits. In each occupational classification a better estimate of Y_{Ai} will be

Table 14. 95 Percent Confidence Limits for Absenteeism Rate for Individual Values of the Independent Variables

Occupational Classification	Actual Absenteeism Rate (Days Per Month)	Predicted (Y_{Ai}) Absenteeism Rate (Days Per Month)	Residual Error (Days per Month)	Confidence Limits
Technologist	0.06	1.92	1.86	0,4.4
Messenger	1.50	1.83	0.33	0,4.1
Laboratory Assistant	0	0.16	0.16	0,0.9

made as each independent variable X_i approaches the mean value of the variable $\overline{X}_i$.

The general methodology followed in computing the confidence limits and the procedure for deriving the C_{ij} values is similar to that shown in Appendix II. A listing of the C_{ij} values for the three occupational classifications listed in Table 14 appears in Appendix III. Similar to that discussed in Chapter IV, equation (30) found in Chapter IV can be used to ascertain the 95 percent confidence limits for sample mean values.

Using equation (31) from Chapter IV, the standard partial regression coefficients were computed and summarized in Table 15. Relative rank of importance of the independent variables is summarized in Table 16.

Table 15. Standard Partial Regression Coefficients with Absenteeism as Dependent Variable

Occupational Classification	Salary ${}^{1}b_1$	Age ${}^{1}b_2$	Dependents ${}^{1}b_3$	Previous Employment ${}^{1}b_4$	Length of Service ${}^{1}b_5$
Technologist	0.8278	1.1159	0.1120	0.5455	0.4099
Messenger	0.7008	1.2715	1.4547	0.4130	0.6514
Laboratory Assistant	1.6813	4.0755	0.4465	2.3038	0.8339

Table 16. Relative Rank of Independent Variable with Absenteeism as Dependent Variable

Occupational Classification	Salary ${}^{1}b_1$	Age ${}^{1}b_2$	Dependents ${}^{1}b_3$	Previous Employment ${}^{1}b_4$	Length of Service ${}^{1}b_5$
Technologist	7.4	9.9	1.0	4.9	3.7
Messenger	1.7	3.1	3.5	1.0	1.6
Laboratory Assistant	3.8	9.1	1.0	5.2	1.9

CHAPTER VI

SUMMARY OF CONCLUSIONS AND DISCUSSION

Results

Mathematical relationships were established for each of the 14 selected occupational classifications using selected quantitative employee variables of salary, age, number of dependents, length of service on previous job, and absenteeism. The existence of regression was found in 10 of the 14 occupational classifications and each of these functional relationships could best be fitted by a multiple linear regression equation. The significant independent variables used in each of the 10 equations were salary, age, length of service on previous job, and absenteeism rate. As developed in Chapter IV, the resulting reduced multiple linear regression equations forecasting length of service are as follows:

For Clerks

$$Y_L = -67.9510 + 0.0245X_1 + 0.0817X_2 - 0.1524X_3 - 0.0804X_4 \tag{18}$$

with the multiple coefficient of correlation R = 0.43 and a standard error of estimate of 32.21 months.

For Porters

$$Y_L = -159.7714 + 0.1103X_1 - 0.0093X_2 - 0.0005X_3 - 0.7853X_4 \tag{19}$$

with the multiple coefficient of correlation R = 0.48 and a standard error of estimate of 10.00 months.

For Orderlies

$$Y_L = 174.8983 + 0.0971X_1 + 0.0912X_2 - 0.0864X_3 - 0.5434X_4 \tag{20}$$

with the multiple coefficient of correlation R = 0.49 and a standard error of estimate of 17.66 months.

For Maids

$$Y_L = -496.9486 + 0.3717X_1 + 0.1371X_2 - 0.1995X_3 - 0.2477X_4 \tag{21}$$

with the multiple coefficient of correlation R = 0.87 and a standard error of estimate of 23.21 months.

For Licensed Practical Nurses - White

$$Y_L = -276.6025 + 0.1149X_1 + 0.0446X_2 - 0.0167X_3 - 0.4612X_4 \tag{22}$$

with the multiple coefficient of correlation $R = 0.82$ and a standard error of estimate of 12.48 months.

For Licensed Practical Nurses - Non-White

$$Y_L = - 938.8269 + 0.3828X_1 + 0.1161X_2 - 0.2929X_3 + 7.1387X_4 \tag{23}$$

with the multiple coefficient of correlation $R = 0.87$ and a standard error of estimate of 18.82 months.

For Secretaries

$$Y_L = - 40.5834 + 0.0129X_1 + 0.0498X_2 + 0.1844X_3 - 3.4818X_4 \tag{24}$$

with the multiple coefficient of correlation $R = 0.56$ and a standard error of estimate of 15.94 months.

For Technologists

$$Y_L = + 12.4416 + 0.0214X_1 - 0.3592X_2 + 0.2496X_3 + 6.8749X_4 \tag{25}$$

with the multiple coefficient of correlation $R = 0.70$ and a standard error of estimate of 17.18 months.

For Messengers

$$Y_L = -262.6905 + 0.1759X_1 + 0.0540X_2 - 0.0716X_3 - 1.6536X_4 \quad (26)$$

with the multiple coefficient of correlation R = 0.89 and a standard error of estimate of 3.44 months.

For Elevator Operators

$$Y_L = 59.1588 + 0.0270X_1 + 0.1031X_2 - 0.0081X_3 - 3.7400X_4 \quad (27)$$

with the multiple coefficient of correlation R = 0.92 and a standard error of estimate of 8.49 months.

Mathematical relationships were established among absenteeism rate and the five quantitative employee variables of salary, age, dependents, length of service on previous job and length of service with hospital. As developed in Chapter V a linear relationship was found to best represent the occupational classifications of Technologists, Messengers and Laboratory Assistants. There was not enough evidence to say regression was present in the other occupational classifications. The resulting multiple linear equations are as follows:

For Technologists

$$Y_A = 0.2742 - 0.0009X_1 + 0.0168X_2 \quad -0.1101X_3 - 0.0170X_4 + 0.0149X_5 \tag{48}$$

with the multiple coefficient of correlation R = 0.64 and a standard error of estimate of 0.79 days per month.

For Messengers

$$Y_A = -\ 49.1046 + 0.0265X_1 + 0.0552X_2 \quad - 1.7509X_3 - 0.0240X_4 - 0.0960X_5 \tag{49}$$

with the multiple coefficient of correlation R = 0.81 and a standard error of estimate of 0.96 days per month.

For Laboratory Assistants

$$Y_A = -\ 10.3103 + 0.0009X_1 - 0.0295X_2 \quad + 0.2329X_3 + 0.1014X_4 + 0.0219X_5 \tag{50}$$

with the multiple coefficient of correlation R = 0.99 and a standard error of estimate of 0.22 days per month.

Conclusions

Equations (18) through (27) and equations (48) through (50) are the models developed in this investigation that can be employed to predict the length of service and absenteeism rate for hospital employees in selected occupational classifications. Equation (29) in Chapter IV can be applied to ascertain the confidence limits for predicted values of single observations for length of service and absenteeism rate. Equation (30) in Chapter IV can be applied to ascertain the confidence limits for predicted values of grouped data for length of service and absenteeism rate.

For the four occupational classifications of Maid, Licensed Practical Nurse - White, Licensed Practical Nurse - Non-White and Messenger for which a prediction model was developed to ascertain expected employee length of service, salary was found to be consistently the most significant independent variable in explaining the variability in employee length of service. Age was ranked second in relative importance in the occupational classifications of Maid, Licensed Practical Nurse - White and Licensed Practical Nurse - Non-White.

For the occupational classifications of Clerk and Elevator Operator, age was found to be the most significant independent variable. Salary and absenteeism rate, respectively, were ranked second in relative importance by these two occupational classifications.

Absenteeism rate was the most significant independent variable found in three occupational classifications of Porter, Secretary and Technologist. Age was found to be second in relative importance for the occupational classifications of Porter and Technologist.

Age was found to be the most important variable in explaining the variability among employee absenteeism for the Technologist and Laboratory Assistant. Length of employment on previous job was ranked second in relative importance.

The most significant variable found in the occupational classification Messenger was dependents. Age ranked second.

Discussion

Total explanation of the variability of hospital employees length of service and absence behavior is still difficult to achieve. However, the methodology followed in this investigation and the results obtained will be of benefit to the hospital manager. He will be able to focus his attention on readily available employee data and to assess their relative importance in order to minimize the effects of excessive employee turnover and absenteeism. Application of regression models for forecasting employee length of service and absenteeism will assist hospital managers not only to ascertain essential information necessary in determining monetary and nonmonetary costs of employee turnover and absenteeism, but also to obtain additional insights into the causes of excessive and avoidable turnover and absenteeism.

Monetary costs associated with labor turnover includes the following: (1) training, (2) added labor -- overtime and staffing to maintain a specified level of service, (3) materials and supplies spoilage, (4) employment office expense, (5) advertising and recruiting expense, (6) pre-job orientation and training, (7) accounting and payroll expense. The total cost per termination has been estimated to range from several hundred dollars to a figure many times larger.[48,49] Although the costs for individual employees vary widely among different institutions, excessive and avoidable employee turnover results in significant expense to the institution.

Excessive absenteeism creates additional monetary costs for the following: (1) added labor -- overtime and part-time replacements to maintain a specified level of service, (2) premiums for paid sick days, (3) payroll and accounting office expense, (4) material and supplies spoilage, and (5) loss of effective production.

Nonmonetary costs of labor turnover and absenteeism include: (1) limitation of planning horizons for personnel staffing patterns, (2) inconsistency of work performance, and (3) production and the problems associated with poor morale.

Reasonably accurate estimates of the monetary and non-monetary costs of excessive turnover reveal the magnitude of the problem. From this data, managers can also determine the minimum length of service necessary to cover expenses incurred in the selection and training of an employee. That is, if the manager wishes to retain an employee a minimum desired length of service can be determined by equating the employee's total recruiting, placement, training and supervising costs incurred by the hospital to a predetermined period of time. This time period could be determined when the time period is viewed as a return on the investment.

Likewise, based on predetermined monetary and non-monetary costs, a desirable level of absenteeism can be sought. The primary objective would be to eliminate or minimize excessive or avoidable absenteeism without penalizing those employees with unavoidable absences.

Aided by the prediction models, managers can assess the effects of alternative administrative policies that seek to increase the length of service of employees and reduce excessive absenteeism.

The effects of a 10 percent increase in salary on length of service for a given employee can be predicted by entering the new salary figure with the other relevant variables into the appropriate occupational prediction model. A comparison of the calculated length of service with the actual length of service yields an estimate of the addition to this employee's length of service attributable to the salary increase. This additional length of service can be viewed as the expected return for the investment of increased salary.

Table 17. Predicted Increase in Length of Service as a Result of a 10 Percent Increase in Average Salary

Occupational Classification	Actual Length of Service (Months)	Predicted Length of Service (Months)	Net Increase in Length of Service (Months)	Percent Increase in Length of Service
Maid	25.63	73.46	47.83	187%
Licensed Practical Nurse-White	24.20	52.34	28.14	116%
Licensed Practical Nurse -Non-White	40.62	133.94	93.32	230%
Messenger	6.50	32.40	25.90	298%

For example, those occupational classifications that indicated that salary was of most relative importance of the four independent variables could expect an increased length of service as a result of a 10 percent increase in salary. Table 17 summarizes the results of an increase of 10 percent in the average salaries paid to those occupational classifications that indicated that salary was of most relative importance. Mean values for the remaining independent variables remained unchanged when placed into the equations.

Similarly, for those occupational classifications that revealed age as the variable of most relative importance, the expected increase in the length of service due to the hiring of only older prospective employees, can be calculated from the prediction models. For example, a policy to raise the mean age by 10 percent would result in an increase in length of service for selected occupational classifications.

Expected results are contained in Table 18. Recognizing that the prediction models were developed using the termination age rather than hiring age poses certain limitations; however, because of the equally high degree of association of both the hiring age and the termination age with length of service, the expected error would be small. Therefore, an increase in the mean termination age used for purposes of the developed prediction models can be interpreted as a similar increase in the hiring age.

Table 18. Predicted Increase in Length of Service Based on a 10 Percent Increase of the Mean Age

Occupational Classification	Actual Length of Service (Months)	Predicted Length of Service (Months)	Net Increase in Length of Service (Months)	Percent Increase in Length of Service
Clerk	18.94	21.77	2.83	14.94%
Elevator Operator	6.50	7.77	1.27	19.54%

The methodology used in predicting the effects of changes in salary and age for selected occupational classifications can be followed for each variable or combination of variables that pertains to a specific situation.

Extrapolation beyond the limits of the original data must at best be made with caution. Nevertheless, the fact that salary and age were found to be consistently of most relative importance, indicates that these two variables deserve primary attention. They can, moreover, be controlled within certain limits by administrative policies.

Salary increases instituted for the purpose of increasing the length of service must be evaluated in terms of existing salary structure within the hospital and the community. One reason that salary is strongly associated with length of service is the lack of a competitive salary structure. This is true not only among occupational classifications in the hospital but also among similar occupational classifications in the community. Increased length of service of an employee may result from achieving an equitable salary structure that is not only consistent with other hospital departments but also with the community's salary level for this occupational classification. Increased length of service of employees as a result of increased wages to achieve a competitive salary scale must be distinguished from the increase in length of service due to a salary scale that exceeds prevailing local and regional scales.

The expected increase in the length of service of an employee paid a salary greater than the "going salary" in the community would require further investigation. Experience of the University of Alabama Hospitals and Clinics has been that increased mobility of certain occupational classifications among community hospitals can be attributed to relatively small changes in salaries, i.e., an employee's move to those institutions that pay a salary slightly higher than their present salary.

The high degree of association between length of service and salary does not necessarily imply that within a particular occupational classification that differences in length of service are caused by differences in salary. In fact the differences in salary may be caused by differences in length of service. However, as used in these forecasting models the implied assumption is that differences in length of service are caused by differences in salary. While the relationship between length of service and salary is found by the higher degree of association between the two variables, the validity of the assumption of cause and effect is subject to further investigation.

Prediction of the effects of changes in the independent variables on the absenteeism rate of employees can also be made by following the methodology developed in this study. Prediction models developed for selected occupational classifications assist in predicting the effects in absenteeism rate which result from changes in the independent variables; however, these changes must also be made with caution. The sign (or slope of line) associated with each partial regression coefficient reflects the relationship each independent variable has with the dependent variable. If a positive relationship exists, the sign of the partial regression coefficient is positive and thus describes the magnitude of the increase in the dependent variable with a unit increase in the independent variables. An evaluation of signs of the partial regression coefficients provides a useful guide in formulating administrative policies that seek to reduce excessive absenteeism.

However, as discussed in the scope and limitations section, the existing sick leave and absenteeism policy often has as much influence on employee absenteeism as any one other variable, e.g., a liberal sick leave policy designed for the conscientious employee may actually encourage or motivate other employees to take sick days needlessly. Nevertheless, the prediction models developed in this investigation can assist hospital management in ascertaining the expected monetary costs such as paid sick days, overtime and relief and, certain nonmonetary costs such as the lack of flexibility in staffing schedules, the variability of workload and problems associated with poor morale. The nature of these nonmonetary costs is difficult to state in explicit terms; consequently the hospital manager must bring to bear on the problem his insights and experiences to realize the full impact of excessive employee absenteeism on hospital activity.

With the knowledge of the expected magnitude and cost of employee absenteeism, the hospital manager can weigh the direct costs with other factors pertaining to employee availability to achieve an acceptable solution to staffing problems. Moreover, the causes can be identified by an analysis of the relationships that are shown in prediction models and an understanding of the effects of the prevailing sick leave policy.

The objective of this investigation was not to identify all causes of turnover; however, in the process of this investigation the following subjective observations of possible causes for turnover were made: (1) inadequate selection and assignment methods, (2) lack of opportunity for advancement; especially in the unskilled occupations, (3) inadequate and poor supervision, (4) inadequate wage structure, (5) inadequate training programs, and (6) inadequate community, especially child care, facilities and convenient transportation.

Observation of possible causes of absenteeism include: (1) poor supervision, (2) inadequate selection, training or promotion programs, (3) lack of child care facilities, and (4) poor transportation.

Knowledge of these possible causes alone is inadequate for the solution of excessive turnover and absenteeism. However, they are important in the formulation of administrative policies designed to minimize excessive turnover and absenteeism. Assisted by the prediction models developed in this investigation, the hospital manager can not only determine nature and magnitude of the problem but also evaluate alternative courses of administrative actions. Thus stated as explicit relationships, the prediction models provide management with an objective method for analyzing the true relationship among selected variables and for evaluating the effects of turnover and absenteeism with changes in selected variables.

Recommendations

This investigation revealed that additional research is specifically needed in three areas related to excessive labor turnover and absenteeism.

1. Analysis of the relationship and effects of differing wage and salary and fringe benefit structures on the mobility of employees among health institutions and facilities in specific geographic areas.
2. Determination of the relationship and effects of other variables such as local transportation systems on employee availability, absenteeism and turnover.
3. Evaluation of financial incentive programs and fringe benefits for retaining employees, improving job advancement, increasing productivity, reducing absenteeism. This evaluation should include the relative values placed upon the programs by different age, sex and occupational classifications.

APPENDIX I

Table 19. Occupational Classifications

Occupational Code	Occupational Title
1	Vice President, Dean
2	Administrator, Assistant Administrator, Administrative Assistant
3	Professor
4	Associate Professor
5	Assistant Professor
6	Instructor
7	Student Assistant
8	Laboratory Consultant
9	Librarian
10	Intern, Extern, Resident
11	Administrative Resident
12	Physician
13	Medical Technologist
14	Technician
15	Laboratory Assistant
16	Technical Aide
17	Embalmer
18	Nurse Anesthetist
19	Chemist
20	Auditor
21	Accountant
22	Collection Manager
23	I.B.M. Operator
24	Cashier
25	Clerical
26	Secretary, Stenographer, Registrars
27	Hostess, Receptionist
28	P.B.X. Operator
29	Director Volunteer Service and Assistant
30	Personnel Director and Assistant
31	Methods Improvement Engineer and Assistant
32	Purchasing Agent and Assistant
33	Storekeeper and Assistant
34	Printer and Assistant
35	Admitting Director and Assistant
36	Dietitian
37	Manager Dietetic Stores

Table 19. Occupational Classifications (continued)

Occupational Code	Occupational Title
38	Cafeteria Manager and Supervisor
39	Assistant Food Service Supervisor
40	Chef, Cook (Chef and Assistant)
41	Housekeeper and Assistant
42	Porter
43	Maid
44	Messenger
45	Linen Supervisors and Checkers
46	Seamstress and Mender
47	Elevator Operator
48	Plant Engineer
49	Maintenance Assistant
50	Security Guard
51	Resident Director
52	Pharmacist
53	Physical Therapist
54	Social Service Worker
55	E.K.G. Director
56	E.E.G. Director
57	Oxygen Therapist
58	Registered Nurse - Administrative
59	Registered Nurse - Supervisor
60	Registered Nurse - Head Nurse (White)
61	Registered Nurse - Head Nurse (Non-White)
62	Registered Nurse - Assistant Head Nurse Team Leader (White)
63	Registered Nurse - Assistant Head Nurse Team Leader (Non-White)
64	Registered Nurse - Staff Nurse (White)
65	Registered Nurse - Staff Nurse (Non-White)
66	Licensed Practical Nurse (White)
67	Licensed Practical Nurse (Non-White)
68	Nursing Aide (White)
69	Nursing Aide (Non-White)
70	Nursing Maid
71	Orderly
72	Professional Nurse Student

Table 19. Occupational Classifications (continued)

Occupational Code	Occupational Title
73	Practical Nurse Student
74	Optical Technician
75	Elevator Supervisor
76	Buyer
77	Administrative Supervisor
78	Autoclave Attendents and Sterile Supply Technicians
79	Venipuncturist
80	Physical Therapist Attendant
81	Nursing Assistant
82	Ward Manager
83	Patient Unit Aide

APPENDIX II

Table 20. Means of Independent and Dependent Variables for Selected Occupational Classifications

Occupational Classification	Salary (Dollars Per Year)	Termination Age (Months)	No. of Dependents	Previous Employment (Months)	Absenteeism (Days Per Month)	Length of Service (Months)
Clerk	2530.21	340.80	0.52	24.04	1.01	18.05
Porter	1563.57	424.23	2.29	106.55	1.14	7.78
Orderly	1703.23	315.87	0.97	54.35	1.77	13.73
Registered Nurse - White	4161.82	321.75	0.41	21.36	0.98	12.02
Maid	1282.89	367.00	1.66	27.82	0.76	24.55
Licensed Practical Nurse - White	2430.97	448.94	0.48	16.00	0.86	21.03
Licensed Practical Nurse- Non-White	2437.50	349.59	2.22	13.66	1.13	39.00
Secretary	3312.33	289.77	0.53	20.27	0.60	18.37
Technologist	5193.64	300.36	0.45	15.86	0.55	23.73
Messenger	1468.24	234.06	0.47	12.18	1.04	5.71
Maintenance Assistant	3876.47	459.29	2.65	37.06	0.32	46.24
Elevator Operator	1575.36	306.43	0.64	37.86	0.63	12.36
Technicians	2890.33	269.50	0.75	17.33	0.26	10.42
Laboratory Assistant	2036.00	257.70	0.20	5.10	0.92	10.10

Table 21. Variances of Independent and Dependent Variables for Selected Occupational Classifications

Occupational Classification	Salary (Dollars Per Year)	Termination Age (Months)	No. of Dependents	Previous Employment (Months)	Absenteeism (Days Per Month)	Length of Service (Months)
Clerk	175808.56	20663.09	1.00	2540.04	5.35	1217.81
Porter	19.64	2782.79	6.43	15136.05	2.84	124.57
Orderly	6759.92	12524.51	3.15	11490.27	20.44	386.46
Registered Nurse - White	38191.97	11092.70	0.71	1389.45	2.10	110.44
Maid	6340.04	17217.95	4.07	2563.94	0.49	1955.88
Licensed Practical Nurse - White	15329.03	18904.93	1.19	456.00	152.19	413.97
Licensed Practical Nurse- Non-White	6212.90	5405.86	3.66	306.75	0.60	1332.32
Secretary	244025.40	2164.25	1.15	837.72	0.54	319.90
Technologist	523405.20	2761.86	0.64	644.89	0.79	468.11
Messenger	652.94	491.06	0.64	274.65	1.83	42.72
Maintenance Assistant	430161.77	18164.97	2.24	2793.43	0.09	1172.69
Elevator Operator	3301.79	22891.65	0.71	11413.21	0.37	335.48
Technicians	407841.33	780.64	0.57	201.33	0.19	57.36
Laboratory Assistant	168960.00	925.79	0.18	24.99	1.40	69.88

Table 22. Standard Error of Estimate for Partial Regression Coefficients with Length of Service as Dependent Variable

Occupational Classification	Sample Size	Salary (Dollars Per Year)	Termination Age (Months)	No. of Dependents	Previous Employment (Months)	Absenteeism (Days Per Month)
Clerk	94	0.0083	0.0269	3.6741	0.0742	1.4976
Porter	91	0.0240	0.0010	0.4753	0.0137	0.6413
Orderly	62	0.0280	0.0378	1.5621	0.0364	0.5086
Registered Nurse - White	44	0.0088	0.0216	2.2488	0.0567	1.2413
Maid	38	0.0461	0.0354	2.0737	0.0793	5.0599
Licensed Practical Nurse - White	31	0.0230	0.0180	2.5791	0.1145	0.1963
Licensed Practical Nurse - Non-White	32	0.0455	0.0519	2.1495	0.2070	4.7035
Secretary	30	0.0077	0.0858	3.2065	0.1491	4.5199
Technologist	22	0.0063	0.1693	5.6037	0.3037	5.0992
Messenger	17	0.0446	0.0901	2.9958	0.0587	1.0446
Maintenance Assistant	17	0.0118	0.0606	6.0331	0.1814	28.7586
Elevator Operator	12	0.6095	0.0298	3.6182	0.3436	3.9515
Technician	12	0.0052	0.1060	4.8783	0.2364	5.9762
Laboratory Assistant	10	0.0110	0.3852	5.6317	1.3309	11.0592

Table 23. Standard Error of Estimate for Partial Regression Coefficients with Absenteeism as Dependent Variable

Occupational Classification	Sample Size	Salary (Dollars Per Year)	Termination Age (Months)	No. of Dependents	Previous Employment (Months)	Length of Service (Months)
Clerk	94	0.0006	0.0020	0.2573	0.0053	0.0076
Porter	91	0.0045	0.0066	0.0797	0.0022	0.0180
Orderly	62	0.0080	0.0102	0.4066	0.0098	0.0344
Registered Nurse - White	44	0.0011	0.0026	0.2922	0.0069	0.0207
Maid	38	0.0027	0.0017	0.0813	0.0029	0.0061
Licensed Practical Nurse - White	31	0.0287	0.0169	2.3748	0.1055	0.1668
Licensed Practical Nurse - Non-White	32	0.0033	0.0022	0.0836	0.0085	0.0075
Secretary	30	0.0003	0.0036	0.1419	0.0062	0.0089
Technologist	22	0.0003	0.0085	0.2708	0.0140	0.0109
Messenger	17	0.0161	0.0177	0.6206	0.0149	0.0760
Maintenance Assistant	17	0.0001	0.0005	0.0458	0.0014	0.0019
Elevator Operator	14	0.0486	0.0035	0.3101	0.0274	0.0266
Technician	12	0.0003	0.0090	0.2931	0.0115	0.0267
Laboratory Assistant	10	0.0002	0.0039	0.2332	0.0183	0.0141

Table 24. Mean Squares Used for Calculating F-Ratios Summarized in Table 2

Occupational Classification	Sum of Squares Due to Regression	Mean Square Due to Regression*	Sum of Squares Deviation About Regression	Mean Square About Regression*	F-Ratio**
Clerk	20990.28	4198.06	92266.46	1048.48	4.00
Porter	2642.33	528.47	8569.28	100.82	5.24
Orderly	5854.11	1170.82	17720.23	316.43	3.70
Registered Nurse - White	208.97	41.79	4540.01	119.47	0.35
Maid	58480.09	11696.02	13887.31	433.98	26.95
Licensed Practical Nurse - White	8365.90	1673.18	4053.07	162.12	10.32
Licensed Practical Nurse - Non-White	31743.93	6348.79	9558.07	367.62	17.27
Secretary	2980.49	596.10	6296.48	262.35	2.27
Technologist	5241.95	1048.39	4588.42	286.78	3.66
Messenger	543.62	108.72	139.91	12.72	8.55
Maintenance Assistant	10528.12	2105.62	8234.94	748.63	2.81
Elevator Operator	3834.07	766.81	527.15	65.89	11.64
Technician	446.36	89.27	184.56	30.76	2.90
Laboratory Assistant	477.41	95.48	151.49	37.87	2.52

*Mean Squares are computed by dividing sums of squares by degrees of freedom.

**The F-Ratio is computed by dividing the sums of squares due to regression by the sums of squares about regression.

Table 25. Data Used for Calculating F-Ratios Summarized in Table 3

Occupational Classification	Sums of Squares Due to Regression	Mean Square Due to Regression	Sums of Squares Within Groups	Error Mean Square Within Groups	F-Ratio
Clerk	88831.51	1432.76	3434.95	132.11	10.84
Porter	6141.04	102.35	2428.23	97.13	1.05
Orderly	17384.98	469.86	335.25	17.64	26.64
Registered Nurse-White	3907.64	169.90	632.37	42.16	4.03
Maid	13205.64	695.03	681.67	52.44	13.25
Licensed Practical Nurse - White	2340.22	146.26	1712.75	190.30	0.77
Licensed Practical Nurse - Non-White	2755.73	183.71	6802.34	681.39	0.27
Secretary	5798.98	362.44	497.50	62.18	5.82
Technologist	2572.82	321.60	2015.30	251.91	1.28
Messenger	14.16	7.08	125.75	13.97	0.51
Maintenance Assistant	3459.94	691.99	4775.00	795.83	0.87
Elevator Operator	105.85	52.93	421.30	70.22	0.75
Technician	29.89	14.95	520.17	130.04	0.11
Laboratory Assistant		Insufficient Data			

Table 26. Data Used for Calculating F-Ratios Summarized in Table 4

Occupational Classification	Sums of Squares Due to Regression	Mean Square Due to Regression	Sums of Squares Within Groups	Error Mean Square Within Groups	F-Ratio
Clerk	21131.45	4226.29	3434.95	132.11	31.99
Porter	2642.34	528.47	2428.23	97.13	5.44
Orderly	5854.11	1170.82	335.25	17.64	66.37
Registered Nurse-White	208.97	41.79	632.37	42.16	0.99
Maid	58480.08	11696.02	681.67	52.44	223.04
Licensed Practical Nurse - White	8366.00	1673.20	1712.75	190.30	8.79
Licensed Practical Nurse - Non-White	26874.65	5374.93	6802.34	681.39	7.89
Secretary	2799.69	559.94	497.50	62.18	9.01
Technologist	5242.24	1048.44	2015.30	251.91	4.16
Messenger	523.62	104.72	125.75	13.97	7.50
Maintenance Assistant	10490.12	209.80	4775.00	795.83	0.26
Elevator Operator	3834.06	766.81	421.30	70.22	10.92
Technician	124.76	24.95	520.17	130.04	0.19
Laboratory Assistant		Insufficient Data			

Table 27. Standard Error of Estimate for Partial Regression Coefficients with Length of Service as Dependent Variable for Selected Occupational Classifications

Occupational Classification	Sample Size	Salary (Dollars Per Year)	Termination Age (Months)	Previous Employment (Months)	Absenteeism (Days Per Month)
Clerk	94	0.0081	0.0252	0.0738	1.4651
Porter	91	0.0238	0.0097	0.0134	0.6383
Orderly	62	0.0277	0.0349	0.0362	0.5046
Registered Nurse - White	44	0.0086	0.0199	0.0566	1.2357
Maid	38	0.0511	0.0338	0.0883	5.6143
Licensed Practical Nurse - White	31	0.0187	0.0176	0.1104	0.1921
Licensed Practical Nurse - Non-White	32	0.0437	0.0465	0.1984	4.4713
Secretary	30	0.0068	0.0774	0.1376	4.4104
Technologist	22	0.0064	0.1527	0.2985	5.1758
Messenger	17	0.0354	0.0395	0.0535	0.6691
Maintenance Assistant	17	0.0124	0.0592	0.1485	2.7597
Elevator Operator	12	0.5595	0.0295	0.3053	4.0479
Technician	12	0.0039	0.0647	0.1677	5.0241
Laboratory Assistant	10	0.0127	0.4467	1.5397	12.7873

Sample Calculations of Confidence Intervals

The normal matrix of the adjusted sum of squares and cross products for the occupational classification of Porter is given below as A:

$$A = \begin{matrix} 176764.28515625 & -11160.00024414 & -28648.57128906 & -297.32304764 \\ -11160.00024414 & 2504510.15625000 & 1397279.45312500 & 27.33821535 \\ -28648.57128906 & 1397279.45312500 & 1362244.53125000 & 2411.62118530 \\ -297.32304764 & 27.33821535 & 2411.62118530 & 255.34927177 \end{matrix}$$

The inverse of matrix A is as follows:

$$A^{-1} = \begin{matrix} 0.00000569 & -0.00000009 & 0.00000020 & 0.00000475 \\ -0.00000009 & 0.00000096 & -0.00000100 & 0.00000923 \\ 0.00000020 & -0.00000100 & 0.00000179 & -0.00001659 \\ 0.00000475 & 0.00000923 & -0.00001659 & 0.00407739 \end{matrix}$$

If $A_i = X_i - \overline{X}_j$, where A_i represents the independent variables then

$$(A_1\ A_2\ A_3\ A_4) \begin{pmatrix} C_{11} & C_{12} & C_{13} & C_{14} \\ C_{21} & C_{22} & C_{23} & C_{24} \\ C_{31} & C_{32} & C_{33} & C_{34} \\ C_{41} & C_{42} & C_{43} & C_{44} \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \end{pmatrix}$$

is equal to $\Sigma\Sigma C_{ij}\ (X_i - \overline{X}_i)\ (X_j - \overline{X}_j)$

The sample calculations for the Porter occupational classification is as follows:

$$(36\ 6\ -17\ -0.14) \quad (A^{-1}) \begin{pmatrix} 36 \\ 6 \\ -17 \\ -\ 0.14 \end{pmatrix} = 0.0087$$

Applying the standard error of estimate of 99.9375, where N = 91, and the values of randomly selected independent variables which are:

X_1	(Salary)	= \$1,564 dollars per year
X_2	(Age)	= 424 months
X_3	(Previous Employment)	= 107 months
X_4	(Absenteeism Rate)	= 1.14 days per month

to the following equation (equation 29 on page 50)

$$Y_L = \hat{Y} \pm t_{0.025}\sqrt{S_e^2 \left(1 + \frac{1}{N} + \Sigma\Sigma C_{ij}\ (X_i-\bar{X}_i)\ (X_j-\bar{X}_j)\right)}$$

the 0.95 percent confidence limits can be obtained.

$\hat{Y} = a + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4$, where a is the mean length of service for the occupational classification, b_i's are the net regression coefficients for the multiple linear regression equation and the X_i's are the values for the independent variables. The values of the independent variable for a given occupational classification are calculated by subtracting the single values of the randomly selected independent variable from the respective mean values (see Table 28) for the independent variable.

Calculation of the predicted $\hat{Y}$ value is as follows:

$$\hat{Y} = 7.78 + 0.1103\ (1560 - 1564) - 0.0093\ (228 - 424)$$
$$- 0.0005\ (0 - 107) - 0.7853\ (0.45 - 1.14)$$
$$= 9.8 \text{ months}$$

Inserting the value of $\hat{Y}$ of 9.8 into equation (29) the 0.95 percent confidence limits are computed as follows:

$$Y_L = 9.8 \pm 1.99\ \sqrt{99.9375\ (1 + \frac{1}{91} - 0.166)}$$

$$Y_L = 9.8 \pm 1.99\ (9.94)$$

$Y_L = 9.8 \pm 19.8$ or an interval of 0 months to 29.6 months.

Inverting the normal matrix (adjusted sum of squares and cross-products) of the independent variables yields the C_{ij} values. With length of service as the dependent variable these C_{ij} values are needed to calculate confidence limits. The C_{ij} values for the occupational classifications used in Chapter IV are as follows:

Clerks

0.00000006	-0.00000000	-0.00000010	0.00000101
-0.00000000	0.00000061	-0.00000068	0.00000153
-0.00000010	-0.00000068	0.00000525	-0.00001650
0.00000101	0.00000153	-0.00001650	0.00206901

Porters

0.00000569	-0.00000009	0.00000020	0.00000475
-0.00000009	0.00000096	-0.00000100	0.00000923
0.00000020	-0.00000100	0.00000179	-0.00001659
0.00000475	0.00000923	-0.00001659	0.00407739

Orderlies

0.00000247	0.00000034	-0.00000019	0.00000195
0.00000034	0.00000391	-0.00000329	-0.00000686
-0.00000019	-0.00000329	0.00000420	0.00000570
0.00000195	-0.00000686	0.00000570	0.00081658

Maids

0.00000485	-0.00000101	0.00000216	0.00001549
-0.00000101	0.00000213	-0.00000260	0.00001633
0.00000216	-0.00000260	0.00001447	-0.00019275
0.00001549	0.00001633	-0.00019275	0.05852527

Licensed Practical Nurses - White

0.00000226	-0.00000019	-0.00000199	0.00000185
-0.00000019	0.00000199	-0.00000243	-0.00000581
-0.00000199	-0.00000243	0.00007828	0.00000327
0.00000185	-0.00000581	0.00000327	0.00023679

Licensed Practical Nurses - Non-White

0.00000541	0.00000001	0.00000364	0.00008413
0.00000001	0.00000610	-0.00000332	-0.00005356
0.00000364	-0.00000332	0.00011113	0.00039140
0.00008413	-0.00005356	0.00039140	0.05645312

Secretaries

0.00000019	0.00000075	-0.00000172	0.00001440
0.00000075	0.00002358	-0.00002317	0.00038171
-0.00000172	-0.00002317	0.00007455	-0.00095430
0.00001440	0.00038171	-0.00095430	0.07652521

Technologists

0.00000014	-0.00000101	0.00000067	0.00006037
-0.00000101	0.00007899	-0.00013169	-0.00100065
0.00000067	-0.00013169	0.00030168	0.00118980
0.00006037	-0.00100065	0.00118980	0.09066416

Messengers

0.00010616	-0.00001777	0.00003069	0.00050497
-0.00001777	0.00013170	-0.00000410	-0.00030600
0.00003069	-0.00000410	0.00024218	0.00059244
0.00050497	-0.00030600	0.00059244	0.03777082

Elevator Operators

-0.00003142	0.00001066	0.00001759	0.00064500
0.00001066	-0.00000003	-0.00000571	0.00001131
0.00001759	-0.00000571	-0.00000280	-0.00009619
0.00064500	0.00001131	-0.00009619	0.21635459

Table 28. Occupational Classification, Mean and Randomly Selected Single Values of Independent Variables Used for Calculating Confidence Limits Summarized in Table 7

OCCUPATIONAL CLASSIFICATION	SALARY (Dollars Per Month) $\bar{X}_1$	X_1	AGE (Months) $\bar{X}_2$	X_2	PREVIOUS EMPLOYMENT (Months) $\bar{X}_3$	X_3	ABSENTEEISM RATE (Days Per Month) $\bar{X}_4$	X_4
Clerk	2530	2400	341	312	24	36	1.01	0
Porter	1564	1560	424	228	107	0	1.14	0.45
Orderly	1703	1670	316	444	54	72	1.77	3.00
Maid	1283	1250	367	396	28	60	0.76	0.90
Licensed Practical Nurse - White	2431	2400	449	264	16	0	3.42	0
Licensed Practical Nurse - Non-White	2438	2520	350	372	14	0	1.13	1.25
Secretary	3312	2820	290	252	20	0	0.60	0.30
Technologist	5194	4860	300	312	16	9	0.55	0.66
Messenger	1468	1500	234	312	12	2	1.04	0.35
Elevator Operator	1575	1560	306	264	38	12	0.63	0

Table 29. Standard Partial Regression Coefficients with Length of Service as Dependent Variable for Selected Occupational Classifications

Occupational Classification	Salary $^1b_1$	Age $^1b_2$	Previous Employment $^1b_4$	Absenteeism Rate $^1b_5$
Clerk	0.3168	0.3636	0.2384	0.0057
Porter	0.0050	0.1551	0.0049	0.1683
Orderly	0.4462	0.3096	0.5217	0.1280
Maid	1.2753	0.7747	0.4347	0.0007
Licensed Practical Nurse - White	1.1383	0.5131	0.2852	0.4562
Licensed Practical Nurse - Non-White	1.6029	0.3299	0.2722	0.2934
Secretary	0.0399	0.1454	0.3344	16.0925
Medical Technologist	0.9003	1.0977	0.3684	35.6154
Messenger	1.3045	0.3475	0.3442	0.6496
Elevator Operator	0.1826	1.8332	0.1005	0.2689

APPENDIX III

Table 30. Mean Squares Used for Calculating F-Ratios Summarized in Table 10

Occupational Classification	Sum of Squares Due to Regression	Error Mean Square Due to Regression*	Sum of Squares Deviation About Regression	Error Mean Square About Regression*	F-Ratio**
Clerk	30.32	6.06	467.46	5.31	1.14
Porter	14.51	2.90	240.83	2.83	1.02
Orderly	48.65	9.73	1198.47	21.40	0.45
Registered Nurse - White	14.50	2.90	75.71	1.99	1.46
Maid	1.14	0.23	16.90	0.53	0.43
Licensed Practical Nurse - White	1120.42	224.08	3445.28	137.81	1.63
Licensed Practical Nurse - Non-White	3.28	0.66	15.31	0.59	1.11
Secretary	3.31	0.66	12.48	0.52	1.27
Technologist	6.78	1.36	9.89	0.62	2.19
Messenger	19.12	3.82	10.18	0.93	4.13
Maintenance Assistant	80.24	0.16	0.57	0.05	3.12
Elevator Operator	1.30	0.26	3.56	0.44	0.58
Technician	1.25	0.25	0.83	0.14	1.82
Laboratory Assistant	12.38	2.48	0.19	0.05	51.43

*Mean Squares are computed by dividing sums of squares by degrees of freedom.

**The F-Ratio is computed by dividing the sums of squares due to regression by the sums of squares about regression.

Table 31. Data Used for Calculating F-Ratios Summarized in Table 11

Occupational Classification	Sum of Squares Due to Regression	Mean Square Due to Regresssion	Sum of Squares Within Groups	Error Mean Square Within Groups	F-Ratio
Clerk		Insufficient Data			
Porter		Insufficient Data			
Orderly		Insufficient Data			
Registered Nurse - White		Insufficient Data			
Maid		Insufficient Data			
Licensed Practical Nurse - White		Insufficient Data			
Licensed Practical Nurse - Non-White	10.91	0.61	4.40	0.55	1.11
Secretary	7.68	0.48	4.80	0.60	0.80
Technologist	9.26	1.16	0.63	0.08	14.50
Messenger	5.89	1.18	4.29	0.72	1.64
Maintenance Assistant	0.41	0.05	0.16	0.05	1.00
Elevator Operator	2.22	0.55	1.34	0.34	1.62
Technician	0.73	0.24	0.10	0.03	8.00
Laboratory Assistant	0.01	0.01	0.18	0.06	0.17

Table 32. Data Used for Calculating F-Ratios Summarized in Table 12

Occupational Classification	Sum of Squares Due to Regression	Mean Square Due to Regression	Sum of Squares Within Groups	Error Mean Square Within Groups	F-Ratio
Clerk		Insufficient Data			
Porter		Insufficient Data			
Orderly		Insufficient Data			
Registered Nurse - White		Insufficient Data			
Maid		Insufficient Data			
Licensed Practical Nurse - White		Insufficient Data			
Licensed Practical Nurse - Non-White	4.89	0.98	4.40	0.55	1.78
Secretary	2.90	0.58	4.80	0.60	0.96
Technologist	6.86	1.37	0.63	0.08	17.13
Messenger	19.16	3.83	4.29	0.72	5.32
Maintenance Assistant	0.80	0.16	0.16	0.05	3.20
Elevator Operator	1.29	0.26	1.34	0.34	0.76
Technician	1.23	0.25	0.10	0.03	8.33
Laboratory Assistant	12.69	2.54	0.18	0.06	42.33

The C_{ij} values for the occupational classifications used in Chapter V, where the dependent variable was absenteeism rate, are as follows:

Technologists

0.00000016	-0.00000170	0.00002902	0.00000080	-0.00000337
0.00000170	0.00011646	-0.00200957	-0.00015730	0.00007376
0.00002902	-0.00200957	0.11864416	0.00180694	-0.00133747
0.00000080	-0.00015730	0.00180694	0.00031816	-0.00004912
-0.00000337	0.00007376	-0.00133747	-0.00004912	0.00019544

Messengers

0.00028223	0.00002088	0.00013095	-0.00001947	-0.00093087
0.00002088	0.00034178	-0.00923432	-0.00000283	0.00038040
0.00013095	-0.00923432	0.41621425	-0.00023048	-0.02567283
-0.00001947	-0.00000283	-0.00023048	0.00024274	0.00022709
-0.00093087	0.00038040	-0.02567283	0.00022709	0.00624257

Laboratory Assistants

0.00000063	-0.00000302	-0.00002222	0.00000553	0.00001189
-0.00000302	0.00032619	-0.01024571	0.00068250	-0.00088877
-0.00002222	-0.01024571	1.12938352	-0.05071132	0.03502607
0.00000553	0.00068250	-0.05071132	0.00702033	-0.00205745
0.00001189	-0.00088877	0.03502607	-0.00205745	0.00410637

LITERATURE CITED

[1]"Statistical Abstract of the United States - 1966," U. S. Department of Commerce, Bureau of the Census, p. 70.

[2]"Guide Issue," Hospitals, J.A.H.A., Vol. 39, No. 15 (August 1, 1965) Part 2, pp. 450-451.

[3]May, Ann. "Absenteeism Survey: Hospitals and Industry," The Modern Hospital, Vol. 90, No. 5 (May, 1958), pp. 87-90, 140.

[4]Levine, E. and Wright, S. "New Ways to Measure Personnel Turnover in Hospitals," Hospitals, J.A.H.A., Vol. 31, (August 1, 1957) pp. 38-42 (September 1, 1957) pp. 50 ff, (October 1, 1957) pp. 47-52.

[5]"Guide Issue," Hospitals, J.A.H.A., Vol. 41, No. 15 (August 1, 1967) Part 2, p. 445.

[6]Weimer, Edward W. "Training Programs Will Help Fill Manpower Gap," Hospitals, J.A.H.A., Vol. 40, No. 11 (June 1, 1966) pp. 83-88, 131.

[7]Crosby, E. L. "Health Manpower in the United States: Review and Reappraisal," Trustee, 18:6 (January, 1965).

[8]Keppel, Francis. "Major Addresses - National Responsibility for Health Manpower," Training Health Service Workers: The Critical Challenge, Proceedings of the Department of Labor-Department of Health, Education and Welfare Conference on Job Development and Training for Workers in Health Services, Washington, D. C., February 14-17, 1966.

[9]Levine and Wright, op. cit.

[10]May, *op. cit.*

[11]Flagle, C. D., Huggins, William H. and Roy, Robert H. *Operations Research and Systems Engineering*. Baltimore, Maryland: Johns Hopkins Press. pp. 763-785.

[12]Lane, K. F. and Andrew, J. E. "A Method of Labour Turnover Analysis," *Journal of the Royal Statistical Society*. Series A (General) 118, Part 3 (1955), pp. 296-323.

[13]Gross, Llewellyn Z., Ph.D.; Yeracoris, Constantine A., Ph.D.; and Grosof, Elliott H., Ph.D. *Social Causes of Labor Turnover in Modern Hospitals*. Report on a Research Project sponsored by the U. S. Public Health Service, National Institutes of Health, Division of Hospital and Medical Facilities (Grant No. HM-00123). Buffalo: Department of Sociology, State University of New York at Buffalo, February, 1966.

[14]Gibson, R. Oliver. "Toward a Conceptualization of Absence Behavior of Personnel in Organizations," *Administrative Science Quarterly* (June, 1966) pp. 107-133. Ithaca, New York: The Graduate School of Business and Public Administration, Cornell University.

[15]Metzner, Helen and Mann, Floyd. "Employee Attitudes and Absences," *Personnel Psychology*, 6 (1953) pp. 467-485.

[16]Shepherd, R. D. and Walker, J. "Absence from Work in Relation to Wage Level and Family Responsibility," *British Journal of Industrial Medicine* (London), 15 (1958), pp. 53 ff.

[17]Mann, Floyd C. and Hoffman, Richard. *Automation and the Worker: A Study of Social Change in Power Plants*. New Ork: Henry Holt, 1960, p. 220.

[18]Metzner and Mann, *op. cit.*

[19]Isambert-Jamati, Viviane. "Absenteeism Among Women Workers in Industry," International Labour Review, 85 (March, 1962), pp. 248-261.

[20]Kahne, Hilda R.; Ryder, Claire F.; Snegriff, Leonid S. and Wyshak, Grace. "Don't Take Older Workers for Granted," Harvard Business Review, 35 (November-December, 1957), pp. 90-94.

[21]Kahne, Ryder, Snegriff and Wyshak, op. cit.

[22]Kahne, Ryder, Snegriff and Wyshak, op. cit.

[23]Gibson, op. cit.

[24]Metzner and Mann, op. cit.

[25]Shepherd and Walker, op. cit.

[26]Acton Society Trust. A Preliminary Study of Attendance at Work in Large and Small Units, and a Further Study of Attendance at Work in Large and Small Units. London: The Acton Society Trust, 1953 and 1957.

[27]Patchen, Martin. "Absence and Employee Feelings About Fair Treatment," Personnel Psychology, 13 (Autumn, 1960) pp. 349-360.

[28]Jewell, Robert E. "An Educational Theory Model: Theory of Local and Cosmopolitan Influentials". Occasional Paper, pp. 63-140. Columbus, Ohio: Bureau of Educational Research and Service, 1963.

[29]Higbie, Edgar Creighton. An Objective Method for Determining Certain Fundamental Principles in Secondary Agricultural Education, published by author, Madison, Wisconsin, 1924.

[30]George, James P. "Job Evaluation Through the Medium of Multiple and Partial Correlation and Regression Analysis," Journal American Statistical Association, Vol. 52, pp. 370-371, September, 1957.

[31]Beenhakker, H. L. "Multiple Correlation - A Technique for Prediction of Future Hospital Bed Needs," <u>Operations Research</u>, Vol. 11, No. 5, September-October 1963, pp. 824-39.

[32]Davis, Edward W. <u>Forecasting the Demand for Hospital Supply Items</u>, M.S. Thesis, Georgia Institute of Technology, January, 1960.

[33]Sonnendecker, J. P. <u>A Model for Forecasting the Whole Blood Requirements of a Hospital Blood Laboratory</u>, M.S. Thesis, Ohio State University, 1958.

[34]May, <u>op. cit.</u>

[35]Groner, Pat N. "A Sick-Leave Policy for the Hospital," <u>The Hospital Digest</u>, Vol. 2, No. 8, pp. 12-13, (August, 1960).

[36]Flagle, Huggins and Roy, <u>op. cit.</u>

[37]Lane and Andrew, <u>op. cit.</u>

[38]Kilbridge, Maurice D. "Turnover, Absence and Transfer Rates as Indicators of Employee Dissatisfaction with Repetitive Work," <u>Industrial and Labor Relations Review</u>, 15 (October, 1961), pp. 21-32.

[39]Gibson, <u>op. cit.</u>

[40]Snedecor, G. W. <u>Statistical Methods</u>, The Iowa State College Press, Ames, Iowa, 1956.

[41]Wine, R. Lowell. <u>Statistics for Scientists and Engineers</u>. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1964.

[42]Ezekiel, M. and Fox, K. A. <u>Methods of Correlation and Regression Analysis</u>, New York: John Wiley & Sons, Inc., 1959, pp. 535.

[43]Ezekiel and Fox, <u>op. cit.</u>, p. 295.

[44]Fisher, R. A. Statistical Methods for Research Workers. New York: Hafner Publishing Company, Inc., 1958, pp. 193-4.

[45]Ezekiel and Fox, op. cit., p. 297.

[46]Ezekiel and Fox, op. cit., p. 285.

[47]Duncan, A. J. Quality Control and Industrial Statistics. Homewood, Illinois: Richard D. Irwin, Inc., 1959, p. 690.

[48]Gaudet, F. J. "Calculating the Cost of Labor Turnover," Personnel, Vol. 35, August 1, 1957, pp. 38-42.

[49]U. S. Employment Service, Government Printing Office. Suggestions for Control of Turnover and Absenteeism, (Rev. Ed. 1957).

OTHER REFERENCES

Flagle, C. D. "Criteria of Effectiveness for Selection, Training and Motivation of Hospital Personnel," (Abstract) Operations Research, Vol. 5, No. 4, August, 1959.

Gaudet, F. J. Solving the Problems of Employee Absence. American Management Association Research Study 57. New York: American Management Association, 1963.

McKenna, G. P. A Mathematical Model for Economic Aggregate Forecasting, M.S. Thesis, Georgia Institute of Technology, August, 1962.

Public Health Service Publication No. 930-C-9, Hospital Personnel, October, 1964.